

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:38:59 ; Search time 48.4 Seconds
(without alignments)
1559,432 Million cell updates/sec

Title: US-09-509-994-1
Perfect score: 2916
Sequence: 1 MGLVYLGLALAGLGRFPAP.....PSPTPGSTLTPPAVGLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

- Database :
- 1: SP-archaea:*
 - 2: SP-bacteria:*
 - 3: SP-fungi:*
 - 4: SP-human:*
 - 5: SP-invertebrate:*
 - 6: SP-mammal:*
 - 7: SP-mhc:*
 - 8: SP-organelle:*
 - 9: SP-phage:*
 - 10: SP-plant:*
 - 11: SP-rodent:*
 - 12: SP-virus:*
 - 13: SP-vertebrate:*
 - 14: SP-unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|--------|--------------------|
| 1 | 2654 | 91.0 | 468 | 4 | Q9UC32 | Q9UC32 homo sapien |
| 2 | 1840 | 63.1 | 577 | 11 | Q35370 | Q35370 rattus norv |
| 3 | 1396.5 | 47.9 | 461 | 11 | P97883 | P97883 rattus norv |
| 4 | 583 | 20.0 | 652 | 4 | Q9NPY3 | Q9NPY3 homo sapien |
| 5 | 580.5 | 19.9 | 757 | 4 | Q9HCU0 | Q9HCU0 homo sapien |
| 6 | 543 | 18.6 | 644 | 11 | Q89103 | Q89103 mus musculu |
| 7 | 530.5 | 18.2 | 643 | 11 | Q9ET61 | Q9ET61 rattus norv |
| 8 | 528.5 | 18.1 | 643 | 11 | Q9J126 | Q9J126 rattus norv |
| 9 | 433 | 14.8 | 1664 | 5 | Q9W021 | Q9W021 caenorhabdi |
| 10 | 414 | 14.2 | 1574 | 11 | Q88281 | Q88281 rattus norv |
| 11 | 385 | 13.2 | 1394 | 5 | Q9V589 | Q9V589 dirosophila |
| 12 | 375 | 12.9 | 708 | 13 | P87363 | P87363 gallus gall |
| 13 | 374 | 12.8 | 2906 | 11 | Q9WUH9 | Q9WUH9 rattus norv |
| 14 | 366 | 12.6 | 3857 | 11 | Q88840 | Q88840 mus musculu |
| 15 | 360 | 12.3 | 528 | 11 | Q9CXDB | Q9CXDB mus musculu |
| 16 | 360 | 12.3 | 2872 | 11 | Q9WUH8 | Q9WUH8 rattus norv |
| 17 | 359 | 12.3 | 741 | 4 | Q9NS21 | Q9NS21 homo sapien |
| 18 | 358.5 | 12.3 | 1174 | 11 | Q99K58 | Q99K58 mus musculu |

| | | | | | | |
|----|-------|------|------|----|--------|--------------------|
| 20 | 353.5 | 12.1 | 1242 | 4 | Q9NS15 | Q9NS15 homo sapien |
| 21 | 353.5 | 12.1 | 1382 | 4 | Q9H7K2 | Q9H7K2 homo sapien |
| 22 | 351.5 | 12.1 | 1511 | 4 | Q75412 | Q75412 homo sapien |
| 23 | 351.5 | 12.1 | 1587 | 4 | Q00508 | Q00508 homo sapien |
| 24 | 341 | 11.7 | 576 | 4 | Q9Y3V7 | Q9Y3V7 homo sapien |
| 25 | 338 | 11.6 | 1062 | 11 | Q60789 | Q60789 mus musculu |
| 26 | 333.5 | 11.4 | 961 | 11 | Q9E0C6 | Q9E0C6 mus musculu |
| 27 | 333 | 11.4 | 999 | 4 | Q9N036 | Q9N036 homo sapien |
| 28 | 331 | 11.4 | 589 | 5 | Q9T2S1 | Q9T2S1 caenorhabdi |
| 29 | 330.5 | 11.3 | 596 | 11 | Q99K64 | Q99K64 mus musculu |
| 30 | 330 | 11.3 | 2189 | 5 | Q9E0F5 | Q9E0F5 elmeria ten |
| 31 | 329.5 | 11.3 | 1964 | 11 | Q35442 | Q35442 mus musculu |
| 32 | 328.5 | 11.3 | 1713 | 11 | Q88349 | Q88349 mus musculu |
| 33 | 328.5 | 11.3 | 1764 | 11 | Q35806 | Q35806 rattus norv |
| 34 | 327.5 | 11.2 | 1095 | 11 | Q60784 | Q60784 mus musculu |
| 35 | 323 | 11.1 | 1253 | 11 | Q61810 | Q61810 mus musculu |
| 36 | 320.5 | 11.0 | 937 | 5 | Q9BLJ1 | Q9BLJ1 clona intes |
| 37 | 320.5 | 11.0 | 2321 | 4 | Q9Y6L8 | Q9Y6L8 homo sapien |
| 38 | 320.5 | 11.0 | 2321 | 4 | Q9W447 | Q9W447 homo sapien |
| 39 | 319.5 | 11.0 | 1833 | 11 | Q08999 | Q08999 mus musculu |
| 40 | 315.5 | 10.8 | 2281 | 4 | Q9UPL3 | Q9UPL3 homo sapien |
| 41 | 315 | 10.8 | 997 | 11 | Q9JFS0 | Q9JFS0 mus musculu |
| 42 | 314.5 | 10.8 | 3507 | 5 | Q23587 | Q23587 caenorhabdi |
| 43 | 313 | 10.7 | 495 | 4 | Q9H805 | Q9H805 homo sapien |
| 44 | 313 | 10.7 | 798 | 5 | Q18026 | Q18026 caenorhabdi |
| 45 | 312.5 | 10.7 | 1821 | 4 | Q14767 | Q14767 homo sapien |

ALIGNMENTS

RESULT 1

Q9UC32 PRELIMINARY: PRT: 468 AA.

AC Q9UC32: 01-MAY-2000 (TREMREL. 13, Created)

DT 01-MAY-2000 (TREMREL. 13, Last sequence update)

DT 01-JUN-2001 (TREMREL. 17, Last annotation update)

DE THROMBOMODULIN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE.

RX MEDLINE=9293792; Pubmed=8390446;

RA Yamamoto S., Mizoguchi T., Tamaki T., Ohkuchi M., Kimura S., Aoki N.;

RT "Urinary thrombomodulin, its isolation and characterization.";

RL J. Biochem. 113:433-440(1993).

DR HSSP: P07204; 12AQ.

DR InterPro: IPR00152; Asx_hydroxyl.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR001881; EGF_Ca.

DR InterPro: IPR001304; lectin_C.

DR InterPro: IPR001491; Thrombomodulin.

DR Pfam: PF00008; EGF_5.

DR Pfam: PF00059; lectin_C_1.

DR PRINTS: PR00907; THROMBOMODULN.

DR SMART: SM0034; CLECT; 1.

DR SMART: SM00181; EGF_6.

DR PROSITE: PS00010; ASX_HYDROXYL; 1.

DR PROSITE: PS00041; C_TYPE_LECTIN_2; 1.

DR PROSITE: PS01186; EGF_2; 2.

DR PROSITE: PS01187; EGF_CA; 1.

RW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation: Repeat.

SQ SEQUENCE 468 AA; 49444 MW; 4BFEB98EFB86A40 CRC64;

Query Match 91.0%; Score 2654; DB 4; Length 468;

Best Local similarity 99.6%; Pred. No. 8 4e-204;

Matches 466; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Q9 19 AAPDPGSGOCVEHDFALYGPATFLMASQICDGLGHIMTVRSSVAADYISLLIIGD 78

```

Db 1 APAEPQSGSCVCHDFALYPGPATFLNLSQICDGLRGLMTVRSSVADVLSLLND 60
QY 79 GGVGRRLMIGLQLPQCGDPRKRLGLRGFWMTGDNNTSYSHMARLDINGAPLCSPLCV 138
Db 61 GGVGRRLMIGLQLPQCGDPRKRLGLRGFWMTGDNNTSYSHMARLDINGAPLCSPLCV 120
QY 139 AVSAEAATVSEPIWEEOCEVKADGFLCEHFPAICRPLAVPAGAAAASVTYCTPRA 198
Db 121 AVSAEAATVSEPIWEEOCEVKADGFLCEHFPAICRPLAVPAGAAAASVTYCTPRA 180
QY 199 ARGADFOALPYGSSAAVAPLGLQIMCTAPGAVOGHMAREAPGAMPCSVENGCEHACNA 258
Db 181 ARGADFOALPYGSSAAVAPLGLQIMCTAPGAVOGHMAREAPGAMPCSVENGCEHACNA 240
QY 259 ITPGAPRCQCPGALQADGSRCTASATQSCNDLCEHFCVNPQPSYSCKETGYRLAA 318
Db 241 ITPGAPRCQCPGALQADGSRCTASATQSCNDLCEHFCVNPQPSYSCKETGYRLAA 300
QY 319 DQHCEDVDDCILBPSPCQRCVNTQGGFECCHCYPNVDLVDGCVAPVPCFRANCETOC 378
Db 301 DQHCEDVDDCILBPSPCQRCVNTQGGFECCHCYPNVDLVDGCVAPVPCFRANCETOC 360
QY 379 QPLNQTSTLCYCAEAFAPIPHEPHRCMFCNQTACPADCDPNTQASCECEGYIIDGFT 438
Db 361 QPLNQTSTLCYCAEAFAPIPHEPHRCMFCNQTACPADCDPNTQASCECEGYIIDGFT 420
QY 439 CTDIDECENGSGFCGVCNHLPGTEPCICGPDALVHRHIGTDCDSCGYVD 486
Db 421 CTDIDECENGSGFCGVCNHLPGTEPCICGPDALVHRHIGTDCDSCGYVD 468

RESULT 2
O35370 PRELIMINARY; PRT; 577 AA.
AC 035370:
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE THROMBOMODULIN.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
RA Wang J., Yao A., Wang J.Y., Hardin J.W., Hauer-Jensen M.;
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-136 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Yao A., Wang J., Hardin J.W., Hauer-Jensen M.;
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF022743; AAB80760.1; -
DR EMBL: AF022742; AAB80923.1; -
DR HSP: P07204; 1FGD.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001304; lectin_c.
DR InterPro: IPR001491; Thrombomodulin.
DR Pfam: PF00008; EGF_5.
DR PRINTS: PR00097; THROMBOMODULN.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00179; EGF_Ca; 2.
DR SMART: SM00001; EGF_like; 3.
DR PROSITE: PS00010; ASX_HYDROXYL; 2.
DR PROSITE: PS50041; C-TYPE-LECTIN_2; 1.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_Ca; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.

```

```

SQ SEQUENCE 577 AA; 61844 MW; 0BE764C8BF18555F CRC64;
Query Match 63.1%; Score 1840; DB 11; Length 577;
Best Local Similarity 63.8%; Pred. No. 8e-159;
Matches 332; Conservative 49; Mismatches 133; Indels 6; Gaps 5;

QY 1 MLGLVIVGALALADLPAPAPFQPGSGSCVCHDFALYPGPATFLNLSQICDGLRGLM 60
Db 1 MLGLVIVGALALADLPAPAPFQPGSGSCVCHDFALYPGPATFLNLSQICDGLRGLM 60
QY 61 TVRSSVADVLSLLNDGCVGRRLMIGLQLPQCGDPRKRLGLRGFWMTGDNNTSYSH 120
Db 61 TVRSSVADVLSLLNDGCVGRRLMIGLQLPQCGDPRKRLGLRGFWMTGDNNTSYSH 118
QY 121 RMARLDINGAPLCSPLCVASAAVAPLGLQIMCTAPGAVOGHMAREAPGAMPCSV 180
Db 119 RMARLDINGAPLCSPLCVASAAVAPLGLQIMCTAPGAVOGHMAREAPGAMPCSV 178
QY 181 EP-GAAAASVTYCTPRAARGADFOALPYGSSAAVAPLGLQIMCTAPGAVOGHMA 239
Db 179 NTRPREGAHISSTYNTPLVSGADFOALPYGSSAAVAPLGLQIMCTAPGAVOGH 238
QY 240 PGAMDCSVENGCEHACNAITPGAPRCQCPGALQADGSRCTASATQSCNDLCEH 299
Db 239 TGAWNCSEVENGCEHYMCNRSANGPRCVCPGSGDLQADGSRCAKAPVQLCNELC 298
QY 300 PDGGSYSCKMETGYRLAAQADGSRCTASATQSCNDLCEHFCVNPQPSYSCKE 359
Db 299 SDVSGYSCKMETGYRLAAQADGSRCTASATQSCNDLCEHFCVNPQPSYSCKE 358
QY 420 NTQASCECEPEGYIIDGFTCTDIDECENGSGFCGVCNHLPGTEPCICGPDAL 479
Db 419 NSPFCQCPREFTLIDECISICTDIDECISQGECLNNECRNLPGSFECICGPD 478
QY 480 CDSGKY--DGDGSGSGEPSPS-PTPGSTLTPPAVGLVHSG 516
Db 479 CDPPLVLEDSGSGSHPSNPTVYSSITVPFARPMHSG 518

RESULT 3
P97883 PRELIMINARY; PRT; 461 AA.
AC P97883:
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE THROMBOMODULIN (FRAGMENT).
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN CAPILLARY;
RA Wang J., Tran N.D., Schreiber S.S., Zlokovic B.V.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: U90121; AAB49723.1; -
DR HSP: P07204; 1FGD.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001491; Thrombomodulin.
DR Pfam: PF00008; EGF_5.
DR PRINTS: PR00097; THROMBOMODULN.
DR SMART: SM00179; EGF_Ca; 2.
DR SMART: SM00001; EGF_like; 3.
DR PROSITE: PS00010; ASX_HYDROXYL; 2.
DR PROSITE: PS01186; EGF_2; 3.

```


RL J. Biol. Chem. 276:7408-7414(2001).
 DR EMBL: AF279142; AAG00867.1; -.
 DR EMBL: AJ295846; CAC34381.1; -.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00059; lectin_c; 1.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00181; EGF; 3.
 DR SMART: SM00179; EGF_Ca; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 DR Signal.
 FT CHAIN 1 17 POTENTIAL.
 FT SIGNAL 18 757 TOMOR ENDOHELIAL MARKER 1.
 SQ SEQUENCE 757 AA; 80859 MW; C96363EALFD8FFA0 CRC64;

Query Match 19.9%; Score 580.5; DB 4; Length 757;
 Best Local Similarity 33.4%; Pred. No. 2.7e-38;
 Matches 176; Conservative 44; Mismatches 192; Indels 115; Gaps 24;

QY 1 MGVLYVGLALALAGL-G-FAPAPAPPGSGCYEHDCFAIYPGATFLNASQICDGLRHL 59
 DB 2 LRLRLAMAAAGPTLGODPWAAPR---AACGSSCYALFRRRTLEAMRACRELGGDL 58
 QY 60 MYRSSVAADVISLLNGDGVGRRRLMIGLOLPGCGDPKRLRGROWVTGDNNTSY 119
 DB 59 ATRTRTEERORVDLVG--AGPASKRLMIGLOLQANOCLOL--PLRGTTWTTGDDDTAF 114
 QY 120 SRMARLDLNGAPLGPICVAVSAEATVPSEPIWEDQCEVKAAGFLCEFHPRATCRPLA 179
 DB 115 TNMAQ--PASGGPCAPRCVLALEAS-----GEHMLEGSCSTLAVDYLQGFEGACPALQ 168
 QY 180 VERGAAAAAVSTIYTPFAKAGDFQALPYGSSAAV---APLGLQIMCTAPRAVAGHNA 236
 DB 169 DEAGQAGPAAV---YTPRFILVSTEFEMLPFGSVAAVCCAGRASLLCYKQPGGVG-WS 224
 QY 237 REAP---GAWDCSVENGCEHAC-NAIPGARPCQAGALADAGSCVATASATQ--CNLD 291
 DB 225 RARPLIGCT-GCSPDNGCCHCEVEVDHVSRCIEGFLADAGSCDEPCQACPEQO 283
 QY 292 CEHFCAVPNPQPGSYSCMCTGTGLAADQ-HRCEDVDDCTLEBSPCRQCVNTQSGFECH 350
 DB 284 CE-----PGPGQYSGCHRLGFRPAEDPHRCVDTDECOI-AGVQOQMCVNYVGGEFCY 336
 QY 351 CYNNYDLVDECEVPVDPFRANCEYQCCPLNFTSLVCAGEGFAPLPHPHRCQMFQNG 410
 DB 337 CSBGHELE-----ADGIS----- 349
 QY 411 TAGPADCP---NTQASCPCPGYIILDGFICTDIDE--CENGFGCSGVCHNLPGT-- 461
 DB 350 -----CSPAGANGAQAQSDGLD-ELLDGDEDEDEDEAMKAFNGMT-----EMGILM 397
 QY 462 FECTCGDSALVRHIGTDCDSKVDGSDSGEPSPPTPGSTLTPP 508
 DB 398 MEPTOPDFALAVRPSRPED-----REPQIPYEPETWPPP 432

RESULT 6
 ID 000274 PRELIMINARY; PRT; 652 AA.
 AC 000274;
 DT 01-JUL-1997 (T-EMBLrel. 04, Created)
 DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)
 DE 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
 DE C1QR.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97199258; PubMed=9047234;
 RA Neomuncine R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;
 RT "cDNA cloning and primary structure analysis of C1QR(P), the human
 C1q/MBL/SPA receptor that mediates enhanced phagocytosis in vitro.";
 RL Immunity 6:119-129(1997).
 DR EMBL: U94533; AAB53110.1; -.
 DR HSSP: P35555; IEMN.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF; 5.
 DR Pfam: PF00059; lectin_c; 1.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00179; EGF_Ca; 3.
 DR SMART: SM00001; EGF-like; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 3.
 DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_Ca; 3.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 SQ SEQUENCE 652 AA; 68576 MW; B7EAB5F5714A775 CRC64;

Query Match 19.9%; Score 580; DB 4; Length 652;
 Best Local Similarity 31.6%; Pred. No. 2.5e-38;
 Matches 180; Conservative 56; Mismatches 210; Indels 124; Gaps 32;

QY 2 LCVLYVGLALALAGL-G-FAPAPAPPGSGCYEHDCFAIYPGATFLNASQICGLRHL 59
 DB 5 MCLLLLLLLLTPQAGGAGATE---AVCVGACVTAHSGKLSAEMAHNQNQNGNL 60
 QY 60 MYRSSVAADV---VISLLNGDGVGRR--RLMIGLOLPGG-GGDPKRLPLRGFWVT 112
 DB 61 ATVKKEERQHYQRLADLRLREALTLARMSKFWIGLOREKGLDPSL--PLGFGFWG 118
 QY 113 GDNNTSYSRWARLDLNGAPLGPICVA--VSAEATVPSE-PWEEQO-----EYKAD 163
 DB 119 GSEDPIYSWMHKELENNSC--ISKRCVSLLDLSQPLRLNRLPKWSEGGCSGSPGSNIE 176
 QY 164 GFICEFHPRATCRPLAV--EPGAAAAVSTIYTPFAKAGDFQALPYGSSAAVAPLGLQ 222
 DB 177 GFVCKCFSEKMGCRPLALGGP-----OYTTTPPTQTSSLEAVPFASANVA----- 224
 QY 223 MC-TAPGAVQGHM--AREAPGAMD-----CSVENGGCEHAC-NAIPGAR 263
 DB 225 -CGEDDKQETQHYFLCKEKAADVDFDMGSSGFLCVSPKYGCNFNNGGCHQDCFGGDSNF 283
 QY 264 RCQCPAGALADAGRSCTASATQSCNDICE--HFCVPPNDPGSTSCMCTGTGLAADQH 321
 DB 284 LCGCRPGEFLDLDLYTC-ASRNPCCSSSPCRGATCVLGP-HGKNYTCRCPOGYLDSSQL 341
 QY 322 RCEDVDCTLEBSPCRQCVNTQSGFECHCYPNYDLVDECEVPVDPFRANCEYQCCPL 381
 DB 342 DCVVDVDEC--QSPCAQCEVNTPGGFRCCEWGY-----EFGGP-----GGACQDV 386
 QY 382 NOTSYLCVCAEFAPLPHPHRCQMFQNCNOTACPADCDPNTQAS--CECPGEGYIL--DDGF 437
 DB 387 DE-----CALGRSP-----CAQGC-TNTDGSFHCSEDEGYVLAGEGGT 423
 QY 438 ICTDIDEC--ENGFGCSGVCHNLPGTFPICGPGSALVRHIGTDCDSKV-----D 486
 DB 424 QCQDVDECVGPGPLCDLSLFCNTQSGFHCGLPGVLAIPN-GVSCSTMPSVSLGPPSPGPD 482
 QY 487 GGDGSGE-----PPSPPTPGSTLTPPA 509
 DB 483 EEDKKEGEGTIVPRAATASPTRGPECTPA 512

RESULT 7
 OX 089103

| ID | PRELIMINARY; | PRT; | 644 AA. |
|----|--|------|---------|
| AC | 089103: | | |
| DT | 01-NOV-1998 (TEMBLrel. 08, Created) | | |
| DT | 01-NOV-1998 (TEMBLrel. 08, last sequence update) | | |
| DT | 01-JUN-2001 (TEMBLrel. 17, last annotation update) | | |
| DE | C10/MBL/SPA RECEPTOR C1QRP. | | |
| GN | LY68 OR C1QRP OR AA4. | | |
| OS | Mus musculus (mouse). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| OX | NCBI_TaxID=10090; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | SPRAIN-129/SVEV. | | |
| RA | Tenner A.J., Kim T.S.; | | |
| RL | "Identification of the mouse genomic DNA for C1QRP."; | | |
| RT | Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases. | | |
| RT | [2] | | |
| RP | SEQUENCE FROM N.A. | | |
| RA | Petrenko O., Lemischka I.R.; | | |
| RL | "Molecular characterization of AA4, an early marker of hematopoietic | | |
| RT | development."; | | |
| RT | Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases. | | |
| RN | [3] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | SPRAIN-129/SV; | | |
| RX | MEDLINE=99359842; PubMed=10430655; | | |
| RA | Norsworthy P.J., Taylor P.R., Walport M.J., Botto M.; | | |
| RL | "Cloning of the mouse homolog of the 126-kDa human C1q/MBL/SP-A | | |
| RT | receptor, C1QRP."; | | |
| RL | Mamm. Genome 10:789-793(1999). | | |
| DR | EMBL: AF074856; AAC63374.1; - | | |
| DR | EMBL: AF081789; AAC62649.1; - | | |
| DR | EMBL: AF099939; AAC47906.1; - | | |
| DR | EMBL: AF099938; AAC47906.1; JOINED. | | |
| DR | HSSP: P07204.1; IEGT. | | |
| DR | MGD: MGI:106664; Ly68. | | |
| DR | InterPro: IPR000152; Asx_hydroxyl. | | |
| DR | InterPro: IPR000561; EGF-like. | | |
| DR | InterPro: IPR001881; EGF.Ca. | | |
| DR | InterPro: IPR001304; lectin.C. | | |
| DR | Pfam: PF00008; EGF. 5. | | |
| DR | Pfam: PF00059; lectin.c; 1. | | |
| DR | SMART: SM00034; C1ECT. 1. | | |
| DR | SMART: SM00179; EGF.Ca; 3. | | |
| DR | SMART: SM00001; EGF-like; 2. | | |
| DR | PROSITE: PS00010; ASX_HYDROXYL. 3. | | |
| DR | PROSITE: PS50041; C_TYPE_LECTIN_2; 1. | | |
| DR | PROSITE: PS01186; EGF. 2; 3. | | |
| DR | PROSITE: PS01187; EGF.Ca; 3. | | |
| KW | Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; | | |
| KW | Receptor; Repeat. | | |
| SO | SEQUENCE 644 AA; 69354 MW; EB4351648BR8635A CRC64; | | |

| | Query Match | Similarity | 18.66 | Score 543 | DB 11 | Length 644 |
|----|-------------|--|-------|-----------------|------------|------------|
| | Best Local | Similarity | 30.08 | Prod. No. 32-35 | | |
| | Matches 167 | Conservative | 55 | Mismatches 211 | Indels 124 | Gaps |
| QY | 5 | LVVLGALAGLGFPAPEPDPGGSCYEHDFCALYPCGPAFLFNLSQICDGLRGHLMTVRS | 64 | | | |
| | | | | | | |
| DB | 9 | LLLGLLGPPMVG--AAADSQ--AVVCEGTALYTHMWKLSAAEQHRCNNGGALVYKS | 64 | | | |
| | | | | | | |
| QY | 65 | SVAA-----DIVSLILNDGGCGVR--RLMIGLQLPPCGSPKRLGPIRGCFQWYTGNDNTS | 118 | | | |
| | | | | | | |
| DB | 65 | EEEAHHVOALTLQLTKTAPLEAKMGKFWITLOREKGCNYHDL--PKRGSPWVGCGEDTA | 123 | | | |
| | | | | | | |
| QY | 119 | YRRMARLDLNGAPLGGRLCYAVNSAAEAIVPSE--PIWEGQCEV-----KADGTCIEHF | 171 | | | |
| | | | | | | |
| DB | 124 | YSNNTKRAKSSGICIFKCYSLIDLSLTPHSLPKHESPCGTPLEAGNNTIEGFLCKFNF | 163 | | | |
| | | | | | | |
| QY | 172 | PATCPRLP--EPGAAAAVSTYTGTPFAAGADFOALPVGSSAAVAPLGIO-----LM | 223 | | | |
| | | | | | | |

[illegible][illegible]

```

Db 64 SEEAHNOVEALADLTKRABSETYIG--KEWIGLQREKGCYTHDL-PMKGFSWVGGE 120
QY 116 NTSYSRMARLDNG--APLCGPLCAVAASAEATVPSE--PIWEEQOCEV-----KADGFL 166
Db 121 DTTYSNMWKAKSSCSISKRCSYSLIDLKPH--PSHLPKWHESPCCGPDAFGNSTIEGL 178
QY 167 CEHFEPATCRPLAV--EFGAAAANVITYGTPPAARGADFOALPVGSSAAVAPLGLQ---- 221
Db 179 CKFNFGKCSPLALGGE-----OLVYTPPQATTSLSLKAVPASVAVNV--CGDEAESK 231
QY 222 ---LMCTAPPAVOGHMAREAP---GAMDCSVENGCEHAC--NAIRGAPRCOPAGAL 273
Db 232 TNYVLCCKETTAGV--FHWSSGFLCVSPKFGCSFNNGCGQDCCFEGGDSFRCGCRPFRL 290
QY 274 QADGRCTASATQSCNDLCEHFVCPNPDPG-----SYSCMCEGYRLAADQHRG 323
Db 291 LDDLVTYCAS-----RNPCSSNPCTGGCMCHSVPLSENYYCHCPRGYOLDSSQVHC 340
QY 324 EDVDDCLLEPSPQRCVNTQGFECCHCYPNYDLVDG--ECVEPVDPFRANCEYQOQPL 381
Db 341 VDIDEC--EDSPQDCEINTPGSFHCEWGYQSSGSKKEACEDVDEC----- 386
QY 382 NOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNTQAS--CECPGTYLD--DGF 437
Db 387 --TAAYSPCAQG-----CT-----NTDGSFYSCCKEGYIMSGDKST 420
QY 438 ICTDIDECENGFCGVCNHLPGTFECICGPDALVHIGTDGDSG-----KV 485
Db 421 QCEDIDEC--LGNPCDTLICINTDGSFRCCGPAFELAPN--GVSCIRGSMFSLPARPQKE 478
QY 486 DGDGSGSGEPSPSPGSS 503
Db 479 DKGDGKESTVPLTEMPGS 496

RESULT 9
Q9J1Z6 PRELIMINARY; PRT; 643 AA.
AC Q9J1Z6:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE C10RP.
GN C10RP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN RP SEQUENCE FROM N.A.
RC STRAIN=MISTER; TISSUE=LUNG;
RA Dean Y.D., McGreal E.P., Akatsu H., Morgan B.P., Gasque P.;
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: A1160978; AAF80402.1;
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PR00008; EGF_5.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00179; EGF_CA; 3.
DR SMART: SM00001; EGF_Like; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 3.
DR PROSITE: PS00041; C_TYPR_LECTIN_2; 1.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_CA; 3.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 643 AA; 68780 MW; F833D742E02A4931 CRC64;

```

Query Match

18.1%; Score 528.5; DB 11; Length 643;

```

Best Local Similarity 29.7%; Pred. No. 3,2e-34;
Matches 166; Conservative 55; Mismatches 210; Indels 127; Gaps 28;

QY 4 VLIALALALAGLPAPAPAPQPGSQCEHDCALYGPATFFLNASQIDGLRHLMTVR 63
Db 8 LLLGLGLQWAGAAASE-----AVYEGTACTAHNGKLSAEQAHRNENGNINATYK 63
QY 64 SVA---DIVISLLN---GDGCVRRRLMIGLDPRECGDPKGLRGFGYTGDN 115
Db 64 SEEAHNOVEALADLTKRABSETYIG--KEWIGLQREKGCYTHDL-PMKGFSWVGGE 120
QY 116 NTSYSRMARLDNG--APLCGPLCAVAASAEATVPSE--PIWEEQOCEV-----KADGFL 166
Db 121 DTTYSNMWKAKSSCSISKRCSYSLIDLKPH--PSHLPKWHESPCCGPDAFGNSTIEGL 178
QY 167 CEHFEPATCRPLAV--EFGAAAANVITYGTPPAARGADFOALPVGSSAAVAPLGLQ---- 221
Db 179 CKFNFGKCSPLALGGE-----OLVYTPPQATTSLSLKAVPASVAVNV--CGDEAESK 231
QY 222 ---LMCTAPPAVOGHMAREAP---GAMDCSVENGCEHAC--NAIRGAPRCOPAGAL 273
Db 232 TNYVLCCKETTAGV--FHWSSGFLCVSPKFGCSFNNGCGQDCCFEGGDSFRCGCRPFRL 290
QY 274 QADGRCTASATQSCNDLCEHFVCPNPDPG-----SYSCMCEGYRLAADQHRG 323
Db 291 LDDLVTYCAS-----RNPCSSNPCTGGCMCHSVPLSENYYCHCPRGYOLDSSQVHC 340
QY 324 EDVDDCLLEPSPQRCVNTQGFECCHCYPNYDLVDG--ECVEPVDPFRANCEYQOQPL 381
Db 341 VDIDEC--EDSPQDCEINTPGSFHCEWGYQSSGSKKEACEDVDEC----- 386
QY 382 NOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNTQAS--CECPGTYLD--DGF 437
Db 387 --TAAYSPCAQG-----CT-----NTDGSFYSCCKEGYIMSGDKST 420
QY 438 ICTDIDECENGFCGVCNHLPGTFECICGPDALVHIGTDGDSG-----KV 485
Db 421 QCEDIDEC--LGNPCDTLICINTDGSFRCCGPAFELAPN--GVSCIRGSMFSLPARPQKE 478
QY 486 DGDGSGSGEPSPSPGSS 503
Db 479 DKGDGKESTVPLTEMPGS 496

RESULT 10
Q9TVQ2 PRELIMINARY; PRT; 1664 AA.
AC Q9TVQ2:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Y6AG10A.7 PROTEIN.
GN Y6AG10A.7
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodertinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN RP SEQUENCE FROM N.A.
RC Mortimore B.J.;
RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 262:2012-2018(1998).
RN RP SEQUENCE FROM N.A.
RA Ainscough R.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: A117206; CAB60454.1; -

```

DR EMBL: AL110498; CAB60454.1; JOINED.
DR EMBL: AL110498; CAB57911.1; -
DR EMBL: AL117206; CAB57911.1; JOINED.
DR HSP: P00736.1APO
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF00008; EGF: 27.
DR PRINTS: PR00011; EGF/LAMININ.
DR SMART: SM00179; EGF_CA: 4.
DR SMART: SM00001; EGF-like: 18.
DR PROSITE: PS00010; ASX_HYDROXYL: 4.
DR PROSITE: PS00022; EGF_1; UNKNOWN_22.
DR PROSITE: PS01186; EGF_2: 24.
DR PROSITE: PS01187; EGF_CA: 3.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat
SQ SEQUENCE 1664 AA: 192797 MW: A69F093B4C705B32 CRC64;

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Agrayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benson P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Bortova D.C., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kamel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., Meperson D.,
 RA Merkulov G., Mikhina N.V., Mobarry C., Morris J., Moshneff A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheffler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA GIBBS R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003558; AAF50538.1; -;
 DR HSSP: P07204; 1ADX.
 DR FlyBase: FBgn0035798; CG7526.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR InterPro: IPR001491; Thrombomodulin.
 DR Pfam: PF00008; EGF_12.
 DR Pfam: PF00084; sush1; 2.
 DR PRINTS: PR00907; THROMBOMODULN.
 DR SMART: SM00032; CCP_2.
 DR SMART: SM00179; EGF_Ca; 9.
 DR SMART: SM00001; EGF_Like; 5.
 DR PROSITE: PS00010; ASX_HYDROXYL; 8.
 DR PROSITE: PS01186; EGF_2; 9.
 DR PROSITE: PS01187; EGF_Ca; 10.
 KM Calcium-binding; EGF-like domain; glycoprotein; Hydroxylation; Repeat.
 FT NON_TER 1394 1394
 SQ SEQUENCE 1394 AA; 152269 MW; CD29380E3162P68A CRC64;

Query Match 13.2%; Score 385; DB 5; Length 1394;
 Best Local Similarity 32.1%; Pred. No. 1.9e-22;
 Matches 99; Conservative 33; Mismatches 110; Indels 66; Gaps 18;

QY 224 CTAPGAVOG-----HWARAPGAW-----DCSVENGCEHCNATPG 261
 DB 563 CLCPFGYALGDNIHTVSLNSFTIDTSSTSPSAHTCLDIDECSLANGNSHPQONRPG 622
 QY 262 APRQCAGALQADGRSCTASATQSC--NDLCEHFCVPPNDOPGSCMCTGYRLAA 318
 DB 623 GFGCACPLGYALSDMRCC--QDIDECIDNSNGQCSQLCL--NQGPGFACACETGFELTP 677
 QY 319 DQHCEDVDCLILPSPRCQVNTGGFEGHCFTPNYLVGFE--CVPEVPDPC--FRAN 373
 DB 678 DFGCADIDECSDQYGNCSIDICINLGHACACERGYELADKLSCLD--VDECAGLISGG 736
 QY 374 CEYQCOPLNPF--SYLCVABEGFAPRPHRCQMFQNCOTACPA--DCDPRNQASC----- 425

DB 737 CSNHC--INKAGTECCGCLGY--ILNDGR-----SCSPALVGGPPGTORSADCCAP 785
 QY 426 -ECPGTYLDDGFTCTDIDEC--ENGFGCSGVCHNLPGTEPCICGP-----DSALVRHI 476
 DB 786 IECNPGTYLSDDKCVDIDECQKNG--CSHRCNTEGSEKSCPPGEYELDSQKTCODI 844
 QY 477 GTDCDSGR 484
 DB 845 -DECDQDK 851
 RESULT 13
 P87363
 ID P87363 PRELIMINARY; PRT; 708 AA.
 AC P87363;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE FTBRILLIN-1 (FRAGMENT).
 GN FBNI.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhou G., Price C., Godfrey M.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U888872; ABA48531.1; -;
 DR HSSP: P07204; 1FGD.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-Like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR002212; TB.
 DR InterPro: IPR001491; Thrombomodulin.
 DR Pfam: PF00008; EGF; 14.
 DR Pfam: PF00683; TB; 2.
 DR PRINTS: PR00907; THROMBOMODULN.
 DR SMART: SM00179; EGF_Ca; 14.
 DR PROSITE: PS00010; ASX_HYDROXYL; 13.
 DR PROSITE: PS01186; EGF_2; 10.
 DR PROSITE: PS01187; EGF_Ca; 13.
 KM Calcium-binding; EGF-like domain; glycoprotein; Hydroxylation; Repeat.
 FT NON_TER 708 708
 SQ SEQUENCE 708 AA; 76163 MW; C247271CIDF73361 CRC64;

Query Match 12.9%; Score 375; DB 13; Length 708;
 Best Local Similarity 31.8%; Pred. No. 6e-22;
 Matches 112; Conservative 26; Mismatches 108; Indels 106; Gaps 20;

QY 244 DCSVENGCCEHCNATPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPPN--- 300
 DB 223 ECSIMNGCCENFTGSESGYSCSKGQFALMPDRHTCT-----DIDE--CEBNPNIC 272
 QY 301 -----DPPGYSQMCCEHGYRLAADQHRCEVDVDCILPSPC--PQRCVNTGGFGFCHYLP 353
 DB 273 DGGQCTNPGEGRCLCYGFMASDEMKTCVDNVEDLHPNCLSGTCENTGSGFICHQDM 332
 QY 354 NYDLVDGE--CVPEVPDPC--FRANCEYOCPLN--QTSYLCVABEGFA-----PIP 398
 DB 333 GYSGKKGTTGCTD--INEEIGAHNCDRAVCTNIPGSKSCSSSWINGIKCTDLDDCS 391
 QY 399 HEPRHCQMFQNOTACPADCPNTQAS--CEPEGYILDDGFTCTDIDE-----CENG- 448
 DB 392 NGTHKCSPH-----ADC--KNITMGSYRCLCKEY--TGDGFTCTDLDECSENTLNLCENGO 442
 QY 449 -----GF-----CS-----GYCHNLPGTEPCICGDSALV 473
 DB 443 CLNAPGGRCECDMGFLPSLDGKACEDIDECSLPNICYGTCCHNLPGLGFRCEVEGYELD 502
 QY 474 RHIG-----TDCDSGR--VDGSGSGSGEPSPPTPSTLTPRAVGLV 513

| Db | 503 | RSGNGCTDVCACADPTCTGTCGTCNATGAGSYTCSCPPD---- | FELNTRTRVGCV | 550 |
|--------|---|---|---------------|-----|
| RESULT | 14 | | | |
| Q9WMH9 | ID | PRELIMINARY; | PRT; 2906 AA. | |
| Q9WMH9 | Q9WMH9 | | | |
| AC | Q9WMH9; | | | |
| DT | 01-NOV-1999 (TREMBlrel. 12, Created) | | | |
| DT | 01-NOV-1999 (TREMBlrel. 12, last sequence update) | | | |
| DT | 01-JUN-2001 (TREMBlrel. 17, last annotation update) | | | |
| DE | FIBRILIN-2. | | | |
| OS | Rattus norvegicus (Rat). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. | | | |
| OX | NCBI_TaxId-10116; | | | |
| RN | [1] | | | |
| RA | SEQUENCE FROM N.A. | | | |
| RA | Yang Q., Ota K., Tian Y., Kumar A., Wada J., Kashihara N., | | | |
| RA | Wallner E.I., Kanwar Y.S.; | | | |
| RT | "Cloning of rat fibrillin-2 cDNA and its role in branching | | | |
| RT | morphogenesis of rat embryonic lung.", | | | |
| RL | Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases. | | | |
| DR | EMBL; AF135060; AAD34439.1; -. | | | |
| DR | HSP; P35555; 1EMN. | | | |
| DR | InterPro: IPR002086; Aldehyde_dehydr. | | | |
| DR | InterPro: IPR000152; Asx_hydroxyl. | | | |
| DR | InterPro: IPR000561; EGF-like. | | | |
| DR | InterPro: IPR001881; EGF_CA. | | | |
| DR | InterPro: IPR001438; EGF_11. | | | |
| DR | InterPro: IPR002212; TB. | | | |
| DR | Pfam; PF00683; EGF; 46. | | | |
| DR | PRINTS; PRO0010; EGFBL00D. | | | |
| DR | SMART; SM00179; EGF_CA; 42. | | | |
| DR | SMART; SM00001; EGF_1like; 4. | | | |
| DR | PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1. | | | |
| DR | PROSITE; PS00010; ASX_HYDROXYL; 43. | | | |
| DR | PROSITE; PS00022; EGF_1; UNKNOWN_2. | | | |
| DR | PROSITE; PS01186; EGF_2; 36. | | | |
| DR | PROSITE; PS01187; EGF_CA; 43. | | | |
| KW | Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat | | | |
| Q9WMH9 | SEQUENCE 2906 AA; 313572 MW; 9EE64E737044EF58 CRC64; | | | |

| | | | | |
|---------------------------|--------|--------------------|------------|--------------|
| Query Match | 12.8%; | Score 374; | DB 11; | Length 2906; |
| Best Local Similarity | 32.9%; | Pred. No. 2.9e-21; | | |
| Matches 102; Conservative | 24; | Mismatches 92; | Indels 92; | Gaps 18 |

| | | | |
|----|------|---|------|
| QY | 244 | DCSEVNGCGCHACNAIIGARCCCPAGAAIOAQRGCTASATQSCUDLCHEHCVPNPD-- | 301 |
| Db | 1237 | ECMIMNGGCDCTQCGNSRGSEGCSSGAYALMPGRSCA-----DIDE--CENNPDC | 1286 |
| QY | 302 | -----QPSYSQMCETGRIADADHRCEVDVDDCLFSPCP--QRCVNTGGFECHCYP | 353 |
| Db | 1287 | DGGGCTNIPETRYKLCIADGFMAMDMKKTCTIDVNECDLNPNICMFGRCEMTKGSFTICHL | 1346 |
| QY | 354 | NYDLVDJGE--CVEYVDPCC--FRANCEYOCOPLN-QTSYLCVCAEGFA-----PIP | 398 |
| Db | 1347 | GYSYKRAKATGCTD--VDCELCGAINCMHMASCLNVPSFSCSCREGVWGNGICIDDLDECA | 1405 |
| QY | 399 | HEPRQCMFCNOTRACPADCDPTQAS--CEPEGYLLDGSFTCTDIDE-----CEMS-- | 448 |
| Db | 1406 | NGTHQCSI-----NAQC--VNTPGSYRCACSGEGF--TGDGFTCSVDVECAENINCEMGO | 1456 |
| QY | 449 | -----GF-----CS-----GYCHNLDPGFEPCICGPDALY | 473 |
| Db | 1457 | CLNVPAYRCECEMGFTPADSDRSQDIDECSSQNIQVGTCTCNLLPGMFHCICIDDGyGLD | 1510 |
| QY | 474 | RHIG--TDCC | 481 |
| Db | 1517 | RTGGHCTDID | 1526 |

| RESULT | 15 | | |
|--------|--|--------------|---------------|
| 088840 | | PRELIMINARY; | PRT; 3857 AA. |
| ID | 088840 | | |
| AC | 088840; | | |
| DT | 01-NOV-1998 (TrEMBLrel. 08, Created) | | |
| DT | 01-MAY-1998 (TrEMBLrel. 08, Last sequence update) | | |
| DT | 01-JUN-2001 (TrEMBLrel. 17, Last annotation update) | | |
| DE | MUTANT FIBRILLIN-1. | | |
| GN | FN1. | | |
| OS | Mus musculus (Mouse). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Murinae; Mus. | | |
| OX | NCBI_TaxID=10090; | | |
| RA | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN-B10.D2; | | |
| RX | MEDLINE=98069008; PubMed=9405934; | | |
| RA | Bona C.A., Mural C., Casares S., Kasturi K., Nishimura H., Honjo T., Matsuda F., | | |
| RT | "Structure of the mutant fibrillin-1 gene in the light skin (TSK) | | |
| RT | mouse."; | | |
| RL | DNA Res. 4:267-271(1997). | | |
| DR | EMBL: AF007248; AAC62317.1; -. | | |
| DR | HSSP: P35555; IAP1 | | |
| DR | InterPro: IPR000152; Asx_hydroxyl. | | |
| DR | InterPro: IPR002557; Chitin_binding. | | |
| DR | InterPro: IPR000561; EGF-like. | | |
| DR | InterPro: IPR001881; EGF-Ca. | | |
| DR | InterPro: IPR002212; TB | | |
| DR | InterPro: IPR000822; znf-C2H2. | | |
| DR | Pfam: PF00008; EGF_64. | | |
| DR | Pfam: PF00683; TB; 12. | | |
| DR | SMART: SM00494; ChEBD2; 2. | | |
| DR | SMART: SM00179; EGF_CA; 60. | | |
| DR | SMART: SM00001; EGF_like_4. | | |
| DR | PROSITE: PS00010; ASX_HYDROXYL; 61. | | |
| DR | PROSITE: PS00022; EGF_1; UNKNOWN_2. | | |
| DR | PROSITE: PS0186; EGF_2; 50. | | |
| DR | PROSITE: PS0187; EGF_CA; 61. | | |
| DR | PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1. | | |
| KW | Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat | | |
| SO | SEQUENCE 3857 AA; 418503 MW; 5BC061B8C527E04C CRC64; | | |

| | | | | |
|-----------------------|------------------|--------------------|-------------|--------------|
| Query Match | 12.6%; | Score 366; | DB 11; | Length 3857; |
| Best Local Similarity | 30.3%; | Pred. No. 1.7e-20; | | |
| Matches 118; | Conservative 28; | Mismatches 132; | Indels 112; | Gaps 21. |

[illegible]

```

Oy      485 -VDGDSGSGEPPSPPTGSLTPPAVGLV 513
          | : | : | | | | |
Db      2488 CVNTPGSYTCDCPD---FELNPTRVGCV 2513

```

Search completed: May 8, 2002, 12:44:48
Job time: 349 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:39:24 ; Search time 17.04 Seconds

(without alignments)
1110.273 Million cell updates/sec

Title: US-09-509-994-1

Sequence: 1 MGVLYGALALAGLGPAP.....PSPFGSTLTPPAVLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|--------------|--------------------|
| 1 | 2912 | 99.9 | 575 | 1 TRBM_HUMAN | P07204 homo sapien |
| 2 | 1839 | 65.1 | 577 | 1 TRBM_MOUSE | P15306 mus musculu |
| 3 | 1092.5 | 37.5 | 356 | 1 TRBM_BOVIN | P06379 bos taurus |
| 4 | 373 | 12.8 | 2907 | 1 FBN2_MOUSE | O61555 mus musculu |
| 5 | 367.5 | 12.6 | 1184 | 1 FBN2_HUMAN | P98095 homo sapien |
| 6 | 367 | 12.6 | 2871 | 1 FBN1_BOVIN | P98133 bos taurus |
| 7 | 366 | 12.6 | 2871 | 1 FBN1_PIG | O91336 sus scrofa |
| 8 | 365 | 12.5 | 2911 | 1 FBN2_HUMAN | P35556 homo sapien |
| 9 | 363 | 12.4 | 2871 | 1 FBN1_HUMAN | P35555 homo sapien |
| 10 | 361 | 12.4 | 2871 | 1 FBN1_MOUSE | O61554 mus musculu |
| 11 | 358 | 12.3 | 956 | 1 MTN2_HUMAN | O00339 homo sapien |
| 12 | 357 | 12.2 | 1221 | 1 FBN2_MOUSE | P37889 mus musculu |
| 13 | 329.5 | 11.3 | 1964 | 1 MTN4_MOUSE | P31955 mus musculu |
| 14 | 328.5 | 11.3 | 956 | 1 MTN2_MOUSE | O08746 mus musculu |
| 15 | 327.5 | 11.2 | 1712 | 1 TGNB_MOUSE | O00918 rattus norv |
| 16 | 324.5 | 11.1 | 712 | 1 FBN1_CAEEL | O77469 caenorhabd |
| 17 | 320 | 11.0 | 1394 | 1 TGNB_HUMAN | P22064 homo sapien |
| 18 | 315.5 | 10.8 | 703 | 1 FBN4_HUMAN | O95667 homo sapien |
| 19 | 313 | 10.7 | 443 | 1 FBN4_MOUSE | P23142 homo sapien |
| 20 | 308.5 | 10.6 | 443 | 1 FBN4_CRIGR | O55058 cricetulus |
| 21 | 308 | 10.6 | 448 | 1 FBN5_HUMAN | O90875 homo sapien |
| 22 | 307 | 10.5 | 705 | 1 FBN1_MOUSE | O08879 mus musculu |
| 23 | 307 | 10.5 | 684 | 1 FBN1_CHICK | O73775 gallus gall |
| 24 | 306.5 | 10.5 | 443 | 1 FBN4_MOUSE | O94739 mus musculu |
| 25 | 302 | 10.4 | 448 | 1 FBN5_MOUSE | O94739 mus musculu |
| 26 | 296 | 10.2 | 448 | 1 FBN5_MOUSE | O94739 mus musculu |
| 27 | 296 | 10.2 | 2531 | 1 NTG1_MOUSE | O01105 mus musculu |
| 28 | 288.5 | 9.9 | 493 | 1 FBN3_HUMAN | O12805 homo sapien |
| 29 | 283.5 | 9.7 | 2437 | 1 NTG1_MOUSE | O01105 mus musculu |
| 30 | 282.5 | 9.7 | 2703 | 1 NTG1_MOUSE | O01105 mus musculu |
| 31 | 280 | 9.6 | 2318 | 1 NTG1_MOUSE | O01105 mus musculu |
| 32 | 277 | 9.5 | 2531 | 1 NTG1_MOUSE | O01105 mus musculu |
| 33 | 276.5 | 9.5 | 493 | 1 FBN3_MOUSE | O01105 mus musculu |

| | | | | | |
|----|-------|-----|------|--------------|--------------------|
| 34 | 271 | 9.3 | 2444 | 1 NTG1_HUMAN | P46531 homo sapien |
| 35 | 269.5 | 9.2 | 1217 | 1 EGF_MOUSE | P01132 mus musculu |
| 36 | 263.5 | 9.0 | 619 | 1 MTN4_HUMAN | O95460 homo sapien |
| 37 | 257.5 | 8.8 | 2524 | 1 NTG1_XENLA | P21783 xenopus lae |
| 38 | 252.5 | 8.7 | 816 | 1 NTL2_RAT | O62918 rattus norv |
| 39 | 252.5 | 8.7 | 835 | 1 CD97_HUMAN | P48960 homo sapien |
| 40 | 251.5 | 8.6 | 816 | 1 NTL2_HUMAN | O99435 homo sapien |
| 41 | 249.5 | 8.6 | 1133 | 1 EGF_RAT | P07522 rattus norv |
| 42 | 248 | 8.5 | 624 | 1 MTN4_MOUSE | O89029 mus musculu |
| 43 | 247 | 8.5 | 816 | 1 NTL2_MOUSE | O61220 mus musculu |
| 44 | 247 | 8.5 | 816 | 1 NTL2_MOUSE | O90827 gallus gall |
| 45 | 246 | 8.4 | 3051 | 1 NTL3_CAEEL | P34576 caenorhabd |

ALIGNMENTS

| | | | | |
|----------|--|-----------|------|---------|
| RESULT 1 | TRBM_HUMAN | STANDARD: | PRT: | 575 AA. |
| AC | P07204: | | | |
| DT | 01-APR-1988 (Rel. 07, Created) | | | |
| DT | 01-FEB-1991 (Rel. 17, Last sequence update) | | | |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) | | | |
| DE | THROMBOMODULIN PRECURSOR (FETOMODULIN) (TM) (CD141 ANTIGEN). | | | |
| GN | THBD OR TRBM. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=88004395; PubMed=2820710; | | | |
| RA | Suzuki K., Kusumoto H., Deyashiki Y., Nishioka J., Maruyama I., | | | |
| RA | Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S., | | | |
| RT | "Structure and expression of human thrombomodulin, a thrombin | | | |
| RT | receptor on endothelium acting as a cofactor for protein C | | | |
| RT | activation."; | | | |
| RL | EMBO J. 6:1891-1897(1987). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=88024950; PubMed=2822087; | | | |
| RA | Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E., | | | |
| RA | "Human thrombomodulin: complete cDNA sequence and chromosome | | | |
| RT | localization of the gene."; | | | |
| RT | Biochemistry 26:4350-4357(1987). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=87317655; PubMed=2819876; | | | |
| RA | Jachman R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D., | | | |
| RT | "Human thrombomodulin gene is intron depleted: nucleic acid sequences | | | |
| RT | of the cDNA and gene predict protein structure and suggest sites of | | | |
| RT | regulatory control."; | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987). | | | |
| RN | [4] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=88227901; PubMed=2836377; | | | |
| RA | Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H., | | | |
| RA | Deyashiki Y., Maruyama I., Suzuki K., | | | |
| RT | "Gene structure of human thrombomodulin, a cofactor for thrombin- | | | |
| RT | catalyzed activation of protein C."; | | | |
| RL | J. Biochem. 103:281-285(1988). | | | |
| RN | [5] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Mathews L., | | | |
| RL | Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases. | | | |
| RN | [6] | | | |
| RP | CARBOHYDRATE-LINKAGE SITE SER-492, AND MUTAGENESIS. | | | |
| RX | MEDLINE=94029900; PubMed=8216207; | | | |
| RA | Gerleitz B., Hassell T., Vlahos C.J., Parkinson J.F., Bang N.U., | | | |
| RA | Grinnell B.W., | | | |
| RT | "Identification of the predominant glycosaminoglycan attachment site | | | |
| RT | in soluble recombinant human thrombomodulin: potential regulation of | | | |

RT functional by glycosyltransferase competition for serine474.";
 RL Biochem. J. 295:131-140(1993).
 RP [7]
 RP STRUCTURE BY NMR OF 389-407.
 RX MEDLINE=96007474; PubMed=7559494;
 RA Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.;
 RT "The structure of a 19-residue fragment from the C-loop of the fourth
 RT epidermal growth factor-like domain of thrombomodulin.";
 RL J. Biol. Chem. 270:23366-23372(1995).
 RN [8]
 RN STRUCTURE BY NMR OF 364-407.
 RX MEDLINE=96100636; PubMed=8528067;
 RA Weininger D.P., Hunter M.J., Komives E.A.;
 RT "Synthesis, activity, and preliminary structure of the fourth
 RT EGF-like domain of thrombomodulin.";
 RL Protein Sci. 4:1683-1695(1995).
 RN [9]
 RN STRUCTURE BY NMR OF 427-444.
 RX MEDLINE=95034791; PubMed=7947766;
 RA Srinivasan J., Hu S., Hirabai R., Zhu Y., Komives E.A., Ni F.;
 RT "Thrombin-bound structure of an EGF subdomain from human
 RT thrombomodulin determined by transferred nuclear Overhauser
 RT effects.";
 RL Biochemistry 33:13553-13560(1994).
 RN [10]
 RN STRUCTURE BY NMR OF 427-444.
 RX MEDLINE=96276211; PubMed=8745396;
 RA Hirabai R., Komives E.A., Ni F.;
 RT "Structural resiliency of an EGF-like subdomain bound to its target
 RT protein, thrombin.";
 RL Protein Sci. 5:195-203(1996).
 RN [11]
 RN STRUCTURE BY NMR OF 405-444.
 RX MEDLINE=98035729; PubMed=9367781;
 RA Sampol Benitez B.A., Hunter M.J., Weininger D.P., Komives E.A.;
 RT "Structure of the fifth EGF-like domain of thrombomodulin: an
 RT EGF-like domain with a novel disulfide-bonding pattern.";
 RL J. Mol. Biol. 273:913-926(1997).
 RN [12]
 RN VARIANT TED TYR-486.
 RX PubMed=7811989;
 RA Oehlin A.-K., Marlar R.A.;
 RT "The first mutation identified in the thrombomodulin gene in a
 RT 45-year-old man presenting with thromboembolic disease.";
 RL Blood 85:330-336(1995).
 RN [13]
 RN VARIANT TED Y-486, AND VARIANTS T-43; A-79; S-495 AND L-501.
 RX PubMed=9198186;
 RA Oehlin A.-K., Norlund L., Marlar R.A.;
 RT "Thrombomodulin gene variations and thromboembolic disease.";
 RL Thromb. Haemost. 78:396-400(1997).
 RN [14]
 RN VARIANT VAL-473.
 RX PubMed=9157575;
 RA Norlund L., Holm J., Zoller B., Oehlin A.-K.;
 RT "A common thrombomodulin amino acid dimorphism is associated with
 RT myocardial infarction.";
 RL Thromb. Haemost. 77:248-251(1997).
 RN [15]
 RN VARIANT THR-43.
 RX PubMed=9843165;
 RA Doggen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,
 RT Stubbs P.J., Manger Cats V., Ireland H.;
 RT "A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr,
 RT and the risk of myocardial infarction in men.";
 RL Thromb. Haemost. 80:743-748(1998).
 RN [16]
 RN VARIANT VAL-473.
 RX PubMed=11245641;
 RA Wu K.K., Aleksic N., Ahn C., Boerwinkle E., Folsom A.R.,
 RT Juneja H.;
 RT "Thrombomodulin Ala45Val polymorphism and risk of coronary heart
 RT disease.";

RL Circulation 103:1386-1389(2001).
 CC -1- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHelial CELL RECEPTOR
 CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS
 CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE
 CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA
 CC SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,
 CC FACTOR VA AND FACTOR VIII, AND THEREBY REDUCES THE AMOUNT OF
 CC THROMBIN GENERATED.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: ENDOTHelial CELLS ARE UNIQUE IN SYNTHESIZING
 CC THROMBOMODULIN.
 CC -1- POLYMORPHISM: VARIATIONS IN THBD ARE ASSOCIATED WITH AN INCREASED
 CC RISK OF DEVELOPING THROMBOEMBOLIC DISEASES (TED).
 CC -1- DISEASE: DEFECTS IN THBD COULD BE THE CAUSE OF INHERITED TED,
 CC ALSO CALLED INHERITED THROMBOPHILIA. PATIENTS WITH TED HAVE
 CC DEFECTS OF THE HAEMOPOIETIC SYSTEM WHICH CREATES A TENDENCY TO THE
 CC OCCURRENCE OF THROMBOSIS. TED PLAY AN IMPORTANT ROLE IN THE
 CC PATHOGENESIS OF VARIOUS CARDIOVASCULAR DISORDERS.
 CC -1- SIMILARITY: CONTRAINS 6 EGF-LIKE DOMAINS.
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD141 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd141.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X05495; CAA29045.1; -
 DR EMBL: M16552; AAB59508.1; -
 DR EMBL: J02973; AAB61175.1; -
 DR EMBL: D00210; BAA00149.1; -
 DR EMBL: AL049651; CAB51954.1; -
 DR PIR: A27073; A27073.
 DR PIR: A28307; A28307.
 DR PIR: A28680; A29680.
 DR PDB: 1EGT; 15-NOV-95.
 DR PDB: 1EGD; 20-JUN-96.
 DR PDB: 1EGE; 20-JUN-96.
 DR PDB: 1TMR; 08-JUN-95.
 DR PDB: 1ZAO; 29-JAN-96.
 DR PDB: 1ADX; 24-DEC-97.
 DR PDB: 2ADX; 24-DEC-97.
 DR GlycoSuiteDB; P07204; -
 DR MIM: 188040; -
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001491; Thmbomoduln.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00008; EGF; 5.
 DR Pfam: PF00059; lectin_c; 1.
 DR PRINTS: PR00907; THROMBOMODULIN.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00179; EGF_CA; 1.
 DR SMART: SM00001; EGF_LIKE; 5.
 DR PROSITE: PS00010; ASX_HYDROXYL; 2.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01187; EGF_CA; 2.
 DR PROSITE: PS00041; C_Type_Lectin_2; 1.
 KW Endothelial cell; Receptor; blood coagulation; Repeat; Transmembrane;
 KW Glycoprotein; Signal; EGF-like domain; Disease mutation; Polymorphism;
 KW 3D-structure.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 575 THROMBOMODULIN.
 FT DOMAIN 22 515 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 516 539 POTENTIAL.
 FT DOMAIN 540 575 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 241 281 EGF-LIKE 1.
 FT DOMAIN 284 324 EGF-LIKE 2.

| | | | | |
|-----------------------|------------------|---------------------|-----------|-------------|
| Query Match | 65.1%; | Score 1839; | DB 1; | Length 577; |
| Best Local Similarity | 65.5%; | Pred. No. 1.1e-117; | | |
| Matches 341; | Conservative 46; | Mismatches 126; | Indels 8; | Gaps 4 |

| ID | TRBM_BOVIN | STANDARD: | PRT: | 356 AA. |
|----|--|-----------|------|---------|
| AC | P06579; | | | |
| DT | 01-JAN-1988 (Rel. 06, Created) | | | |
| DT | 01-JAN-1988 (Rel. 06, Last sequence update) | | | |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) | | | |
| DE | THROMBOMODULIN (FETOMODULIN) (TM) (FRAGMENT). | | | |
| GN | THBD. | | | |
| OS | Bos taurus (Bovine). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; | | | |
| OC | Bovidae; Bovinae; Bos. | | | |
| OX | NCBI_TaxID=9913; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=87067408; PubMed=3024152; | | | |
| RA | Jackman R.W., Beeler D.L., Vandewater L., Rosenberg R.D.; | | | |
| RT | "Characterization of a thrombomodulin cDNA reveals structural | | | |
| RT | similarity to the low density lipoprotein receptor"; | | | |
| TL | Proc. Natl. Acad. Sci. U.S.A. 83:8834-8838(1986). | | | |
| CC | -1- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR | | | |
| CC | THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS | | | |
| CC | COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE | | | |
| CC | ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA | | | |
| CC | SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM, | | | |
| CC | FACTOR VA AND FACTOR VIIIa. AND THEREBY REDUCES THE AMOUNT OF | | | |

| | | |
|----|--|---|
| CC | | THROMBIN GENERATED. |
| CC | -1- | SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. |
| CC | -1- | TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING THROMBOMODULIN. |
| CC | -1- | SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS. |
| CC | | ----- |
| CC | | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). |
| CC | | ----- |
| DR | EMBL; M14657; AAA30785.1; -. . | |
| DR | PIR; A25918; A25918. | |
| DR | HSSP; P07204; 1EGF. | |
| DR | InterPro; IPRO00152; Asx_hydroxyl. | |
| DR | InterPro; IPR000561; EGF-like. | |
| DR | InterPro; IPR01881; EGF_Ca. | |
| DR | Pfam; PF00008; EGF_5. | |
| DR | SMART; SM00179; EGF_CA; 1. | |
| DR | SMART; SM00001; EGF_Like; 3. | |
| DR | PROSITE; PS00010; ASX_HYDROXYL; 2. | |
| DR | PROSITE; PS00022; EGF_1; FALSE_NNG. | |
| DR | PROSITE; PS01187; EGF_2; 3. | |
| KW | Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane; Glycoprotein; EGF-like domain. | |
| FT | NON_TER | 1 |
| FT | DOMAIN | 1 296 |
| FT | TRANSMEM | 297 320 |
| FT | DOMAIN | 321 356 |
| FT | DOMAIN | 17 57 |
| FT | DOMAIN | 60 98 |
| FT | DOMAIN | 99 137 |
| FT | DOMAIN | 138 178 |
| FT | DOMAIN | 179 214 |
| FT | DOMAIN | 215 254 |
| FT | DISULEID | 21 32 |
| FT | DISULEID | 28 41 |
| FT | DISULEID | 21 32 |
| FT | DISULEID | 43 56 |
| FT | DISULEID | 64 72 |
| FT | DISULEID | 68 82 |
| FT | DISULEID | 84 97 |
| FT | DISULEID | 103 114 |
| FT | DISULEID | 110 123 |
| FT | DISULEID | 125 136 |
| FT | DISULEID | 143 152 |
| FT | DISULEID | 148 162 |
| FT | DISULEID | 164 178 |
| FT | DISULEID | 182 191 |
| FT | DISULEID | 187 199 |
| FT | DISULEID | 201 213 |
| FT | DISULEID | 219 228 |
| FT | DISULEID | 224 237 |
| FT | DISULEID | 239 253 |
| FT | CARBONYD | 271 |
| QO | SEQUENCE | 356 AA; 37795 MW; 29B4LF097ABE1093 CRC64; SIMILARITY). EXTRACELLULAR (POTENTIAL). CYTOPLASMIC (POTENTIAL). EGF-LIKE 1. EGF-LIKE 2. EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL). EGF-LIKE 4. EGF-LIKE 5. EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL). BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. O-LINKED (GLYCOSAMINOGLYCAN) (BY SIMILARITY). |

| | | | | |
|-----------------------|-----------------|-------------------|-----------|------------|
| Query Match | 37.5% | Score 1092.5 | DB 1 | Length 356 |
| Best Local Similarity | 63.4% | Pred. No. 5,1e-65 | | |
| Matches 189 | Conservative 30 | Mismatches 64 | Indels 15 | Gaps 6 |

| | | | |
|----|-----|---|-----|
| QY | 229 | GAVGGHAKAPAPAMDCSYENGGCGHACNALPGAPRCQCPAGALDADAGSCASTQSC | 288 |
| | | | |
| Db | 5 | GEGGGRSRAPAMACGVERGGCGHEKGSAGASNCICPPADALADGSSCLPHEMPC | 64 |
| | | | |
| QY | 289 | NDICEHFCVNPDPGSSYSCMCETGYRLAADONHCEVDCLLEPSPCPQCVNTGGFE | 348 |
| | | | |
| Db | 65 | HQCEHCEHC - HILHGAGNYTCICEAGTCGLAADONHCEVDCLDCLPSPCPQCVNTGGGQ | 122 |
| | | | |

| | | | | |
|----|----------|------|------|----------------|
| FT | DISULFID | 329 | 343 | BY SIMILARITY. |
| FT | DISULFID | 345 | 358 | BY SIMILARITY. |
| FT | DISULFID | 491 | 503 | BY SIMILARITY. |
| FT | DISULFID | 498 | 512 | BY SIMILARITY. |
| FT | DISULFID | 514 | 525 | BY SIMILARITY. |
| FT | DISULFID | 532 | 542 | BY SIMILARITY. |
| FT | DISULFID | 537 | 551 | BY SIMILARITY. |
| FT | DISULFID | 553 | 566 | BY SIMILARITY. |
| FT | DISULFID | 572 | 584 | BY SIMILARITY. |
| FT | DISULFID | 579 | 593 | BY SIMILARITY. |
| FT | DISULFID | 595 | 608 | BY SIMILARITY. |
| FT | DISULFID | 614 | 625 | BY SIMILARITY. |
| FT | DISULFID | 620 | 634 | BY SIMILARITY. |
| FT | DISULFID | 636 | 649 | BY SIMILARITY. |
| FT | DISULFID | 655 | 666 | BY SIMILARITY. |
| FT | DISULFID | 661 | 675 | BY SIMILARITY. |
| FT | DISULFID | 677 | 690 | BY SIMILARITY. |
| FT | DISULFID | 765 | 777 | BY SIMILARITY. |
| FT | DISULFID | 772 | 786 | BY SIMILARITY. |
| FT | DISULFID | 788 | 801 | BY SIMILARITY. |
| FT | DISULFID | 807 | 819 | BY SIMILARITY. |
| FT | DISULFID | 814 | 828 | BY SIMILARITY. |
| FT | DISULFID | 830 | 843 | BY SIMILARITY. |
| FT | DISULFID | 849 | 859 | BY SIMILARITY. |
| FT | DISULFID | 854 | 868 | BY SIMILARITY. |
| FT | DISULFID | 870 | 883 | BY SIMILARITY. |
| FT | DISULFID | 952 | 964 | BY SIMILARITY. |
| FT | DISULFID | 959 | 973 | BY SIMILARITY. |
| FT | DISULFID | 975 | 988 | BY SIMILARITY. |
| FT | DISULFID | 1070 | 1082 | BY SIMILARITY. |
| FT | DISULFID | 1077 | 1091 | BY SIMILARITY. |
| FT | DISULFID | 1093 | 1106 | BY SIMILARITY. |
| FT | DISULFID | 1112 | 1124 | BY SIMILARITY. |
| FT | DISULFID | 1119 | 1133 | BY SIMILARITY. |
| FT | DISULFID | 1135 | 1149 | BY SIMILARITY. |
| FT | DISULFID | 1155 | 1167 | BY SIMILARITY. |
| FT | DISULFID | 1162 | 1176 | BY SIMILARITY. |
| FT | DISULFID | 1178 | 1191 | BY SIMILARITY. |
| FT | DISULFID | 1197 | 1209 | BY SIMILARITY. |
| FT | DISULFID | 1204 | 1218 | BY SIMILARITY. |
| FT | DISULFID | 1220 | 1233 | BY SIMILARITY. |
| FT | DISULFID | 1239 | 1250 | BY SIMILARITY. |
| FT | DISULFID | 1246 | 1259 | BY SIMILARITY. |
| FT | DISULFID | 1261 | 1274 | BY SIMILARITY. |
| FT | DISULFID | 1280 | 1292 | BY SIMILARITY. |
| FT | DISULFID | 1287 | 1301 | BY SIMILARITY. |
| FT | DISULFID | 1303 | 1316 | BY SIMILARITY. |
| FT | DISULFID | 1322 | 1334 | BY SIMILARITY. |
| FT | DISULFID | 1329 | 1343 | BY SIMILARITY. |
| FT | DISULFID | 1345 | 1358 | BY SIMILARITY. |
| FT | DISULFID | 1364 | 1377 | BY SIMILARITY. |
| FT | DISULFID | 1371 | 1386 | BY SIMILARITY. |
| FT | DISULFID | 1386 | 1399 | BY SIMILARITY. |
| FT | DISULFID | 1405 | 1418 | BY SIMILARITY. |
| FT | DISULFID | 1412 | 1427 | BY SIMILARITY. |
| FT | DISULFID | 1429 | 1440 | BY SIMILARITY. |
| FT | DISULFID | 1446 | 1458 | BY SIMILARITY. |
| FT | DISULFID | 1453 | 1467 | BY SIMILARITY. |
| FT | DISULFID | 1469 | 1482 | BY SIMILARITY. |
| FT | DISULFID | 1488 | 1499 | BY SIMILARITY. |
| FT | DISULFID | 1494 | 1508 | BY SIMILARITY. |
| FT | DISULFID | 1510 | 1523 | BY SIMILARITY. |
| FT | DISULFID | 1529 | 1540 | BY SIMILARITY. |
| FT | DISULFID | 1535 | 1549 | BY SIMILARITY. |
| FT | DISULFID | 1551 | 1564 | BY SIMILARITY. |
| FT | DISULFID | 1647 | 1659 | BY SIMILARITY. |
| FT | DISULFID | 1654 | 1668 | BY SIMILARITY. |
| FT | DISULFID | 1670 | 1683 | BY SIMILARITY. |
| FT | DISULFID | 1689 | 1701 | BY SIMILARITY. |
| FT | DISULFID | 1696 | 1710 | BY SIMILARITY. |

Query Match 12.8%; Score 373; DB 1; Length 2907;
 Best Local Similarity 32.9%; Pred. No. 4, 5e-17;

| | | | |
|--|--|---|---------------|
| Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18; | | | |
| QY | 244 | DCSEVNGCGHAONALPGAPROCCPAGALADGADSCGASATQSCNDCEHPCVPPND-- | 301 |
| Db | 1238 | ECMIMGCGCTQCTNBSGSECSGEGVALMPDGSCA-----DIDE--CENMPDIC | 1287 |
| QY | 302 | -----QPSYSQMCETGYRLADQRCEDVDDEILPSPCP-ORCVNTOGGECHCYP | 353 |
| Db | 1288 | DGGCCTNPEYRCLCYDGFPMASMDKKTCIDVNEEDLNPNICMFCEGNTGSGFICHOQL | 1347 |
| QY | 354 | NYDLVGE--CVPVPPC--FRANCYQCPPLN-QTSYLCVCAEFA-----PIP | 398 |
| Db | 1348 | GYSVKKGTTCCTD-VDECEIGAHNCMDHASCINVPSCSCSGEGWNGIKCIDLDECA | 1406 |
| QY | 399 | HEPHRCMFQNCYACPADCDPNTQAS--CECEGYTLDDGFCTDIDE-----CENG- | 448 |
| Db | 1407 | NGTHGCSI-----NAQC-VNTPGSGYRCACSEGF-TGGGFCGSDYDCANTNCENGQ | 1457 |
| QY | 449 | -----GF-----GS-----GVCNLPGETFEICGPDLSLV | 473 |
| Db | 1458 | CLNVPGARCECEMGFTPASDSRSCODIDECGFONICVGTGCTNLPGMFHCICDDGYELD | 1517 |
| QY | 474 | RHIG--TDGD 481 | |
| Db | 1518 | RTGSGNCTDID 1527 | |
| RESULT 5 | | | |
| ID | FBUL2_HUMAN | STANDARD; | PRT: 1184 AA. |
| AC | P88055; | | |
| DT | 01-NOV-1995 (Rel. 32, Created) | | |
| DT | 01-NOV-1995 (Rel. 32, Last sequence update) | | |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) | | |
| DE | FBULIN-2 PRECURSOR. | | |
| GN | FBULN2. | | |
| OS | Homo sapiens (human). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| OX | NCBI_TaxId=9606; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | TISSUE=Brain; Pubmed=7806230; | | |
| RX | MDLINE=95104855; Pubmed=7806230; | | |
| RA | Zhang R.-Z., Pan T.-C., Zhang Z.-Y., Mattei M.-G., Timpl R., | | |
| RA | Chu M.-L., | | |
| RT | "Fibulin-2 (FBULN2): human cDNA sequence, mRNA expression, and mapping | | |
| RT | of the gene on human and mouse chromosomes."; | | |
| RL | Genomics 22:425-430(1994). | | |
| CC | -1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS | | |
| CC | CALCIUM DEPENDENT. | | |
| CC | -1- SUBUNIT: HOMODIMER: DISULFIDE-LINKED. | | |
| CC | -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX. COMPONENT OF BOTH | | |
| CC | BASEMENT MEMBRANES AND OTHER CONNECTIVE TISSUES. | | |
| CC | -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, PLACENTA AND OVARY. | | |
| CC | -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS. | | |
| CC | -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS. | | |
| CC | ----- | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | |
| CC | use by non-profit institutions as long as its content is in no way | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | |
| CC | or send an email to license@sib-sib.ch). | | |
| CC | ----- | | |
| DR | EMBL; X82494; CA57876.1; - | | |
| DR | HSSP; P07204; IFGD. | | |
| DR | MM; 135821; - | | |
| DR | InterPro; IPR000020; Anaphylatoxin. | | |
| DR | InterPro; IPR000152; Asx hydroxyl. | | |
| DR | InterPro; IPR000561; EGF-like. | | |
| DR | InterPro; IPR001881; EGF_Ca. | | |

DR Pfam: PF01821; ANATO; 2.
 DR Pfam: PF00008; EGF; 7.
 DR SMART: SM00104; ANATO; 3.
 DR SMART: SM00179; EGF_CA; 9.
 DR PROSITE: PS00001; EGF_Like; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 5.
 DR PROSITE: PS01177; ANAPHYLATOXIN_1; 3.
 DR PROSITE: PS01178; ANAPHYLATOXIN_2; 3.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 5.
 DR PROSITE: PS01187; EGF_CA; 9.
 DR Signal: Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;
 KW Calcium-binding; Repeat.
 FT CHAIN 1 27
 FT SIGNAL 1 27
 FT DOMAIN 28 1184
 FT DOMAIN 28 444
 FT DOMAIN 178 177
 FT DOMAIN 445 444
 FT DOMAIN 488 519
 FT DOMAIN 521 553
 FT DOMAIN 604 645
 FT DOMAIN 679 718
 FT DOMAIN 719 763
 FT DOMAIN 764 809
 FT DOMAIN 810 857
 FT DOMAIN 858 900
 FT DOMAIN 901 942
 FT DOMAIN 943 981
 FT DOMAIN 982 1024
 FT DOMAIN 1025 1069
 FT DOMAIN 1070 1184
 FT DISULEID 445 472
 FT DISULEID 446 479
 FT DISULEID 459 480
 FT DISULEID 489 518
 FT DISULEID 502 519
 FT DISULEID 521 545
 FT DISULEID 522 552
 FT DISULEID 538 553
 FT DISULEID 608 620
 FT DISULEID 616 629
 FT DISULEID 631 644
 FT DISULEID 683 693
 FT DISULEID 689 702
 FT DISULEID 704 717
 FT DISULEID 723 736
 FT DISULEID 730 745
 FT DISULEID 751 762
 FT DISULEID 768 781
 FT DISULEID 775 790
 FT DISULEID 796 808
 FT DISULEID 814 827
 FT DISULEID 821 836
 FT DISULEID 843 856
 FT DISULEID 862 875
 FT DISULEID 869 884
 FT DISULEID 886 899
 FT DISULEID 905 917
 FT DISULEID 913 926
 FT DISULEID 928 941
 FT DISULEID 947 956
 FT DISULEID 952 965
 FT DISULEID 967 980
 FT DISULEID 986 998
 FT DISULEID 994 1007
 FT DISULEID 1009 1023
 FT DISULEID 1029 1042
 FT DISULEID 1036 1051
 FT DISULEID 1056 1068
 FT CARBOHYD 180 507
 FT CARBOHYD 507 1035

SQ SEQUENCE 1184 AA: 126543 MW: CA48490A55F9EC5D CRC64:
 Query Match 12.6%; Score 367.5; DB 1; Length 1184;
 Best Local Similarity 30.0%; Pred. No. 4.7e-17;
 Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;
 DB 154 EEOCEYKAD-GFLCEHPATC-----RPLAV-----EPGAAAAAVSITVGTPEARG 201
 DB 531 EGQCESNPMLGIPCN-HVHLSCEBGEPLIVEVRPPRPAARVR-----EADM 582
 DB 202 ADFQALPVSSAAV---APRLQIMCTAPGAVQGMAREAPGAWDCSVENGCEHACNA 258
 DB 583 AGRBALSLGTEAEPLNSLPDQDDDECLLPGEI-----CQHLCIN 622
 DB 259 IPGAPRCQCPAGALQADGRSC-----TASAT-----QS 287
 DB 623 TVGSYHCACPGSLQDDGRTCPREGHPQPEAPQEDALKSEFSQVASTIPLPLPQPT 682
 DB 288 CND-LCEHFCVNPDPGSGYSCMETGYRLADQRCEDVDCCILEPSCP--QRCVNT 343
 DB 683 CKDNGPCQKVC---STVGSALSCSPFGVAINADGVSCEDINCYVDLHTSCGECVNT 739
 DB 344 QGFECH---CYPNDLVGCEVEPYDPCFRANCEYOCPLNOSTYLCV----- 389
 DB 740 LGSFHCYKALICEPGYALKDGEV-EDVDEC--AMGTHTCOP---GFLQNTKGSFYCOA 792
 DB 390 ---CAEGFAPLPH-----EPHRCQHECNOTACPADCDNTQASCECPGY-I 432
 DB 793 RQRCMDGFLQDPENCVDINECTSLSEPCRPSCITVTSYTCQNPPLI---CANGYNA 849
 DB 433 LDDGFICTDIDCENGSGCSG---VCHNDPTEECIC 466
 DB 850 SDDGAKCVDVNECETGVHRCGEGOVCHNLDGSRXDC 886
 RESULT 6
 FBNI_BOVIN STANDARD; PRT; 2871 AA.
 ID FBNI_BOVIN
 AC P98133;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE FIBRILLIN 1 PRECURSOR (MP940).
 GN FBNI.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RC MEDLINE=95137597; PubMed=7835900;
 RA Tlstira D.J., Potter K.A., Byers P.H.;
 RT "Sequence of the coding region of the bovine fibrillin cDNA and
 RT localization to bovine chromosome 10.";
 RL Genomics 23:480-485(1994).
 RN [2]
 RN PARTIAL SEQUENCE.
 RX MEDLINE=96132851; PubMed=8557636;
 RA Gibson M.A., Hatzinikolas G., Kumaratilake J.S., Sandberg L.B.,
 RA Nicholl J.K., Sutherland G.R., Cleary E.G.;
 RT "Further characterization of proteins associated with elastic fiber
 RT microfibrils including the molecular cloning of MGP-2 (MP25).";
 RL J. Biol. Chem. 271:1096-1103(1996).
 CC -I- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
 CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
 CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
 CC -I- PM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
 CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
 CC MICROFIBRILS.
 CC -I- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING

```

CC      EGF-LIKE DOMAINS.
CC      -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; L28748; AAT4122.1; -.
DR      HSSP; P35555; IAPJ.
DR      InterPro: IPR000152; Asx_hydroxyl.
DR      InterPro: IPR000561; EGF-like.
DR      InterPro: IPR001881; EGF_Ca.
DR      InterPro: IPR001438; EGF_II.
DR      InterPro: IPR002212; TB.
DR      Pfam; PF00008; EGF; 46.
DR      Pfam; PF00683; TB; 9.
DR      PRINTS; PR00010; EGFBL00D.
DR      SMART; SM00179; EGF_CA; 42.
DR      SMART; SM00001; EGF_Like; 4.
DR      PROSITE; PS00010; ASX_HYDROXYL; 43.
DR      PROSITE; PS00022; EGF_1; 2.
DR      PROSITE; PS01186; EGF_2; 38.
DR      PROSITE; PS01187; EGF_CA; 45.
KW      Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW      Repeat; Signal; Multigene family.
FT      SIGNAL 1 27
FT      CHAIN 28 2871
FT      DOMAIN 81 112 EGF-Like 1, NON-CALCIUM BINDING.
FT      DOMAIN 115 146 EGF-Like 2, NON-CALCIUM BINDING.
FT      DOMAIN 147 178 EGF-Like 3, NON-CALCIUM BINDING.
FT      DOMAIN 246 287 EGF-Like 4, CALCIUM-BINDING.
FT      DOMAIN 288 329 EGF-Like 5, CALCIUM-BINDING.
FT      REPEAT 330 390 TGFBP 1.
FT      DOMAIN 392 446 PRO-RICH.
FT      DOMAIN 449 489 EGF-Like 6, NON-CALCIUM BINDING.
FT      DOMAIN 490 529 EGF-Like 7, CALCIUM-BINDING.
FT      DOMAIN 530 571 EGF-Like 8, CALCIUM-BINDING.
FT      DOMAIN 572 612 EGF-Like 9, CALCIUM-BINDING.
FT      DOMAIN 613 653 EGF-Like 10, CALCIUM-BINDING.
FT      REPEAT 654 722 TGFBP 2.
FT      DOMAIN 723 764 EGF-Like 11, CALCIUM-BINDING.
FT      DOMAIN 765 806 EGF-Like 12, CALCIUM-BINDING.
FT      DOMAIN 807 846 EGF-Like 13, CALCIUM-BINDING.
FT      DOMAIN 910 951 EGF-Like 14, CALCIUM-BINDING.
FT      REPEAT 952 1027 TGFBP 3.
FT      DOMAIN 1028 1069 EGF-Like 15, CALCIUM-BINDING.
FT      DOMAIN 1070 1112 EGF-Like 16, CALCIUM-BINDING.
FT      DOMAIN 1113 1154 EGF-Like 17, CALCIUM-BINDING.
FT      DOMAIN 1155 1196 EGF-Like 18, CALCIUM-BINDING.
FT      DOMAIN 1197 1237 EGF-Like 19, CALCIUM-BINDING.
FT      DOMAIN 1238 1279 EGF-Like 20, CALCIUM-BINDING.
FT      DOMAIN 1280 1321 EGF-Like 21, CALCIUM-BINDING.
FT      DOMAIN 1322 1362 EGF-Like 22, CALCIUM-BINDING.
FT      DOMAIN 1363 1403 EGF-Like 23, CALCIUM-BINDING.
FT      DOMAIN 1404 1445 EGF-Like 24, CALCIUM-BINDING.
FT      DOMAIN 1446 1486 EGF-Like 25, CALCIUM-BINDING.
FT      DOMAIN 1487 1527 EGF-Like 26, CALCIUM-BINDING.
FT      REPEAT 1528 1605 TGFBP 4.
FT      DOMAIN 1606 1647 EGF-Like 27, CALCIUM-BINDING.
FT      DOMAIN 1648 1688 EGF-Like 28, CALCIUM-BINDING.
FT      REPEAT 1689 1765 TGFBP 5.
FT      DOMAIN 1766 1807 EGF-Like 29, CALCIUM-BINDING.
FT      DOMAIN 1808 1848 EGF-Like 30, CALCIUM-BINDING.
FT      DOMAIN 1849 1890 EGF-Like 31, CALCIUM-BINDING.
FT      DOMAIN 1891 1929 EGF-Like 32, CALCIUM-BINDING.
FT      DOMAIN 1930 1972 EGF-Like 33, CALCIUM-BINDING.
FT      DOMAIN 1973 2012 EGF-Like 34, CALCIUM-BINDING.
FT      DOMAIN 2013 2054 EGF-Like 35, CALCIUM-BINDING.
FT      REPEAT 2055 2126 TGFBP 6.

```

```

FT      DOMAIN 2127 2165 EGF-Like 36, CALCIUM-BINDING.
FT      DOMAIN 2166 2205 EGF-Like 37, CALCIUM-BINDING.
FT      DOMAIN 2206 2246 EGF-Like 38, CALCIUM-BINDING.
FT      DOMAIN 2247 2290 EGF-Like 39, CALCIUM-BINDING.
FT      REPEAT 2291 2332 EGF-Like 40, CALCIUM-BINDING.
FT      DOMAIN 2333 2402 TGFBP 7.
FT      DOMAIN 2403 2443 EGF-Like 41, CALCIUM-BINDING.
FT      DOMAIN 2444 2484 EGF-Like 42, CALCIUM-BINDING.
FT      DOMAIN 2485 2523 EGF-Like 43, CALCIUM-BINDING.
FT      DOMAIN 2524 2566 EGF-Like 44, CALCIUM-BINDING.
FT      DOMAIN 2567 2606 EGF-Like 45, CALCIUM-BINDING.
FT      DOMAIN 2607 2647 EGF-Like 46, CALCIUM-BINDING.
FT      DOMAIN 2648 2687 EGF-Like 47, CALCIUM-BINDING.
FT      DISULFID 85 94 BY SIMILARITY.
FT      DISULFID 89 100 BY SIMILARITY.
FT      DISULFID 102 111 BY SIMILARITY.
FT      DISULFID 113 129 BY SIMILARITY.
FT      DISULFID 123 134 BY SIMILARITY.
FT      DISULFID 136 145 BY SIMILARITY.
FT      DISULFID 150 160 BY SIMILARITY.
FT      DISULFID 154 166 BY SIMILARITY.
FT      DISULFID 168 177 BY SIMILARITY.
FT      DISULFID 180 262 BY SIMILARITY.
FT      DISULFID 250 271 BY SIMILARITY.
FT      DISULFID 257 286 BY SIMILARITY.
FT      DISULFID 273 304 BY SIMILARITY.
FT      DISULFID 292 313 BY SIMILARITY.
FT      DISULFID 299 328 BY SIMILARITY.
FT      DISULFID 315 328 BY SIMILARITY.
FT      DISULFID 453 465 BY SIMILARITY.
FT      DISULFID 460 474 BY SIMILARITY.
FT      DISULFID 476 488 BY SIMILARITY.
FT      DISULFID 494 504 BY SIMILARITY.
FT      DISULFID 499 513 BY SIMILARITY.
FT      DISULFID 515 528 BY SIMILARITY.
FT      DISULFID 534 546 BY SIMILARITY.
FT      DISULFID 541 555 BY SIMILARITY.
FT      DISULFID 557 570 BY SIMILARITY.
FT      DISULFID 576 587 BY SIMILARITY.
FT      DISULFID 582 596 BY SIMILARITY.
FT      DISULFID 598 611 BY SIMILARITY.
FT      DISULFID 617 628 BY SIMILARITY.
FT      DISULFID 623 637 BY SIMILARITY.
FT      DISULFID 639 652 BY SIMILARITY.
FT      DISULFID 727 739 BY SIMILARITY.
FT      DISULFID 734 748 BY SIMILARITY.
FT      DISULFID 750 763 BY SIMILARITY.
FT      DISULFID 769 781 BY SIMILARITY.
FT      DISULFID 776 790 BY SIMILARITY.
FT      DISULFID 792 805 BY SIMILARITY.
FT      DISULFID 811 821 BY SIMILARITY.
FT      DISULFID 816 830 BY SIMILARITY.
FT      DISULFID 832 845 BY SIMILARITY.
FT      DISULFID 914 926 BY SIMILARITY.
FT      DISULFID 921 935 BY SIMILARITY.
FT      DISULFID 937 950 BY SIMILARITY.
FT      DISULFID 1032 1044 BY SIMILARITY.
FT      DISULFID 1039 1053 BY SIMILARITY.
FT      DISULFID 1055 1068 BY SIMILARITY.
FT      DISULFID 1074 1086 BY SIMILARITY.
FT      DISULFID 1081 1095 BY SIMILARITY.
FT      DISULFID 1097 1111 BY SIMILARITY.
FT      DISULFID 1117 1129 BY SIMILARITY.
FT      DISULFID 1124 1138 BY SIMILARITY.
FT      DISULFID 1140 1153 BY SIMILARITY.
FT      DISULFID 1159 1171 BY SIMILARITY.
FT      DISULFID 1166 1180 BY SIMILARITY.
FT      DISULFID 1182 1195 BY SIMILARITY.
FT      DISULFID 1201 1212 BY SIMILARITY.
FT      DISULFID 1208 1221 BY SIMILARITY.
FT      DISULFID 1223 1236 BY SIMILARITY.
FT      DISULFID 1242 1254 BY SIMILARITY.
FT      DISULFID 1249 1263 BY SIMILARITY.
FT      DISULFID 1265 1278 BY SIMILARITY.

```


| | | | | | |
|---------------------------|---|---|------|----------|--|
| FT | DISULFID | 1223 | 1236 | | BY SIMILARITY. |
| FT | DISULFID | 1242 | 1254 | | BY SIMILARITY. |
| FT | DISULFID | 1248 | 1263 | | BY SIMILARITY. |
| FT | DISULFID | 1265 | 1278 | | BY SIMILARITY. |
| FT | DISULFID | 1284 | 1296 | | BY SIMILARITY. |
| FT | DISULFID | 1291 | 1305 | | BY SIMILARITY. |
| FT | DISULFID | 1307 | 1320 | | BY SIMILARITY. |
| FT | DISULFID | 1326 | 1339 | | BY SIMILARITY. |
| FT | DISULFID | 1333 | 1348 | | BY SIMILARITY. |
| FT | DISULFID | 1350 | 1361 | | BY SIMILARITY. |
| FT | DISULFID | 1367 | 1380 | | BY SIMILARITY. |
| FT | DISULFID | 1374 | 1389 | | BY SIMILARITY. |
| FT | DISULFID | 1401 | 1402 | | BY SIMILARITY. |
| FT | DISULFID | 1415 | 1429 | | BY SIMILARITY. |
| FT | DISULFID | 1450 | 1461 | | BY SIMILARITY. |
| FT | DISULFID | 1456 | 1470 | | BY SIMILARITY. |
| FT | DISULFID | 1472 | 1485 | | BY SIMILARITY. |
| FT | DISULFID | 1491 | 1502 | | BY SIMILARITY. |
| FT | DISULFID | 1497 | 1511 | | BY SIMILARITY. |
| FT | DISULFID | 1513 | 1526 | | BY SIMILARITY. |
| FT | DISULFID | 1610 | 1622 | | BY SIMILARITY. |
| FT | DISULFID | 1617 | 1631 | | BY SIMILARITY. |
| FT | DISULFID | 1633 | 1646 | | BY SIMILARITY. |
| FT | DISULFID | 1652 | 1663 | | BY SIMILARITY. |
| FT | DISULFID | 1658 | 1672 | | BY SIMILARITY. |
| FT | DISULFID | 1674 | 1687 | | BY SIMILARITY. |
| FT | DISULFID | 1770 | 1782 | | BY SIMILARITY. |
| FT | DISULFID | 1777 | 1791 | | BY SIMILARITY. |
| FT | DISULFID | 1793 | 1806 | | BY SIMILARITY. |
| FT | DISULFID | 1812 | 1824 | | BY SIMILARITY. |
| FT | DISULFID | 1818 | 1833 | | BY SIMILARITY. |
| Query Match | | | | | 12.6%; Score 366; DB 1; Length 2871; |
| Best Local Similarity | | | | | 30.6%; Pred. No. 1.3e-16; |
| Matches 109; Conservative | | | | | 26; Mismatches 107; Indels 114; Gaps 20; |
| OY | 244 | DCSEVNGCGCHACNAIPGAPCCCPAGAAADAGSGTASATQSCNDLCEHFCVNP---- | 300 | | |
| Db | 1200 | ECSIMNCGCTPFTNSEGSYCSCPFGALMPDROSC-----DIDE--CEDPNITC | 1249 | | |
| OY | 301 | -DOPSYSCMETGYRLAODHREVDPCILEFPSPC-PQRCVNTQGFECHCYR | 353 | | |
| Db | 1250 | DGGCGTNIEPEHYKCLCYDFGNASEDMKTICVYNEDDLNPNICLSGTGENTKSGFI | 1309 | | |
| OY | 354 | NYDLVGE--CVEVPDPC--FRANCETOCPLNOT-STLYCABGFA-----PIP | 398 | | |
| Db | 1310 | GYSKKRKTKTCGD-INDECEIGAHNCDRNAVCTNTGSEFNSCSGWMIGIDIKCTDDECS | 1368 | | |
| OY | 399 | HEPRRCOMPONOTACPACDDPTQAS--CEEGEYLIDDFICTDIDEC-EN----- | 447 | | |
| Db | 1369 | NGTAMCQH-----ADC-KNTMSGRKLCKEGT-TDGDTGCADDDECSBNKLCGNVO | 1419 | | |
| OY | 448 | ---GGF-----CS-----GVCHNLPGTFECTICGPSALV | 473 | | |
| Db | 1420 | CLYAPAGHCEYDMGFVPSADRKSVDSDCSLPNICVFCTGHMLPGLFRCEC----- | 1472 | | |
| OY | 474 | RHIGTDSDSKYVGGDSGSEPPS-----PTFGS-----ILTTPAVGLV | 513 | | |
| Db | 1473 | -EIYEELDRSGNCTDVNECLEPTTCISGNCVNTPGSYSTVCVCPPEFLNPTRVGCV | 1527 | | |
| <hr/> | | | | | |
| RESULT | 8 | | | | |
| FBN2_HUMAN | P35556; | STANDARD; | PRT; | 2911 AA. | |
| DT | 01-JUN-1994 (Rel. 29, Created) | | | | |
| DT | 01-FEB-1996 (Rel. 33, Last sequence update) | | | | |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) | | | | |
| DE | FIBRILLIN 2 PRECURSOR. | | | | |
| GN | FBN2. | | | | |
| OS | Homo sapiens (human). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |

FT DISULFID 197 206 BY SIMILARITY.
 FT DISULFID 279 281 BY SIMILARITY.
 FT DISULFID 286 300 BY SIMILARITY.
 FT DISULFID 302 315 BY SIMILARITY.
 FT DISULFID 321 333 BY SIMILARITY.
 FT DISULFID 328 342 BY SIMILARITY.
 FT DISULFID 344 357 BY SIMILARITY.
 FT DISULFID 487 509 BY SIMILARITY.
 FT DISULFID 504 518 BY SIMILARITY.
 FT DISULFID 520 532 BY SIMILARITY.
 FT DISULFID 538 548 BY SIMILARITY.
 FT DISULFID 543 557 BY SIMILARITY.
 FT DISULFID 559 572 BY SIMILARITY.
 FT DISULFID 578 590 BY SIMILARITY.
 FT DISULFID 585 599 BY SIMILARITY.
 FT DISULFID 601 614 BY SIMILARITY.
 FT DISULFID 620 631 BY SIMILARITY.
 FT DISULFID 626 640 BY SIMILARITY.
 FT DISULFID 642 655 BY SIMILARITY.
 FT DISULFID 661 672 BY SIMILARITY.
 FT DISULFID 667 681 BY SIMILARITY.
 FT DISULFID 683 696 BY SIMILARITY.
 FT DISULFID 771 783 BY SIMILARITY.
 FT DISULFID 778 792 BY SIMILARITY.
 FT DISULFID 794 807 BY SIMILARITY.
 FT DISULFID 813 825 BY SIMILARITY.
 FT DISULFID 820 834 BY SIMILARITY.
 FT DISULFID 836 849 BY SIMILARITY.
 FT DISULFID 855 865 BY SIMILARITY.
 FT DISULFID 860 874 BY SIMILARITY.
 FT DISULFID 876 889 BY SIMILARITY.
 FT DISULFID 958 970 BY SIMILARITY.
 FT DISULFID 965 979 BY SIMILARITY.
 FT DISULFID 981 994 BY SIMILARITY.
 FT DISULFID 1076 1088 BY SIMILARITY.
 FT DISULFID 1083 1097 BY SIMILARITY.
 FT DISULFID 1099 1112 BY SIMILARITY.
 FT DISULFID 1118 1130 BY SIMILARITY.
 FT DISULFID 1125 1139 BY SIMILARITY.
 FT DISULFID 1141 1155 BY SIMILARITY.
 FT DISULFID 1161 1173 BY SIMILARITY.
 FT DISULFID 1168 1182 BY SIMILARITY.
 FT DISULFID 1184 1197 BY SIMILARITY.
 FT DISULFID 1203 1215 BY SIMILARITY.
 FT DISULFID 1210 1224 BY SIMILARITY.

Query Match 12.5% Score 365; DB 1; Length 2911;
 Best Local Similarity 33.1%; Pred. No. 1.5e-16;
 Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps 17;

OY 244 DCSVENGECEHACNAIPGAPROCCPAGALQADGSRSCATASATQSCNDLCHEFCVNPDP-- 301
 DB 1244 ECMINMGGDYQCTSESGYSCSGEVALMPDGRSCA-----DIDE--CENNPDIC 1293
 OY 302 -----OPGSYCMCEGTGYRLAADOHRCEDVDCILERSPCP-QRCVNTGGGFECHCYP 353
 DB 1294 DGGQCTNINTEGRCLCYGFASMSMKITIDVNECDLSNIMCEPCENTKGSFICHCOL 1353
 OY 354 NYDLVDGE--CYPEVDPG--FRANCEYOQOPLN-QTSLVCVCAEGFAPRPH-----EPHR 403
 DB 1354 GYSVKKGTGCTD-VDECEIGAHNCDMHASCINIGSRKSCREGW--IGNIKCIDLDE 1410
 OY 404 COMFCNOTACPADCDPNTQAS--CECPBGYIILDDGFICTIDDE-----CENG----- 448
 DB 1411 CNGTHQCSINAGC-VNTPGYSRACSEGF--TGDEFTSDVDECAENINLCEMGCOLNVP 1468
 OY 449 -----GF-----CSGVCHNLPGTEFCICGPPSALVRHIG- 477
 DB 1469 GAYRECECEGFTPADSRSCQDIDRCSFONICVSGTCNMLPMGFHCTDDGYELDRGTGN 1528
 OY 478 -TDCD 481
 DB 1529 CTDDID 1533

RESULT 9
 ID FBN1 HUMAN STANDARD; PRT; 2871 AA.
 AC P35555;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE FIBRILLIN 1 PRECURSOR.
 GN FBN1 OR FBN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=93372860; PubMed=8364578;
 RA Pereira L.V., D'Alessio M., Ramirez F., Lynch J.R., Sykes B.,
 Pangillan T., Bonadio J.;
 RT "Genomic organization of the sequence coding for fibrillin, the
 defective gene product in Marfan syndrome.";
 RL Hum. Mol. Genet. 2:961-968(1993).
 RN [2]
 RP SEQUENCE OF 1-932 FROM N.A.
 RC TISSUE=Placenta; and Fibroblast;
 RX MEDLINE=94010947; PubMed=7691719;
 RA Corson G.M., Chabberg S.C., Dietz H.C., Charbonneau N.L., Sakai L.Y.;
 RT "Fibrillin binds calcium and is coded by cDNAs that reveal a
 multidomain structure and alternatively spliced exons at the 5'
 end.";
 RL Genomics 17:476-484(1993).
 RN [3]
 RP SEQUENCE OF 899-2871 FROM N.A.
 RX MEDLINE=91304567; PubMed=1852207;
 RA Maslen C.L., Corson G.M., Maddox B.K., Gianvillie R.W., Sakai L.Y.;
 RT "Partial sequence of a candidate gene for the Marfan syndrome.";
 RL Nature 352:334-337(1991).
 RN [4]
 RP SEQUENCE OF 813-1313 FROM N.A.
 RX MEDLINE=91304567; PubMed=1852206;
 RA Lee B., Godfrey M., Vitale E., Hort H., Mattei M.-G., Sartarzi M.,
 Tsipouras P., Ramirez F., Hollister D.W.;
 RT "Linkage of Marfan syndrome and a phenotypically related disorder to
 two different fibrillin genes.";
 RL Nature 352:330-334(1991).
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE=91317849; PubMed=1860873;
 RA Sakai L.Y., Keene D.R., Gianvillie R.W., Bachinger H.P.;
 RT "Purification and partial characterization of fibrillin, a cysteine-
 rich structural component of connective tissue microfibrils.";
 RL J. Biol. Chem. 266:14763-14770(1991).
 RN [6]
 RP STRUCTURE BY NMR OF 2054-2125.
 RX MEDLINE=98031893; PubMed=9362480;
 RA Yuan X., Downing A.K., Knott V., Handford P.A.;
 RT "Solution structure of the transforming growth factor beta-binding
 protein-like module, a domain associated with matrix fibrils.";
 RL EMBO J. 16:6659-6666(1997).
 RN [7]
 RP STRUCTURE BY NMR OF 2124-2205.
 RX MEDLINE=96144829; PubMed=8568869;
 RA Knott V., Downing A.K., Cardy C.M., Handford P.A.;
 RT "Calcium binding properties of an epidermal growth factor-like domain
 pair from human fibrillin-1.";
 RL J. Mol. Biol. 255:22-27(1996).
 RN [8]
 RP STRUCTURE BY NMR OF 2124-2205.
 RX MEDLINE=96222301; PubMed=8653794;
 RA Downing A.K., Knott V., Werner J.M., Cardy C.M., Campbell I.D.,
 RA Handford P.A.;

RT "Solitistic structure of a pair of calcium-binding epidermal growth
RT factor-like domains: implications for the Marfan syndrome and other
RT genetic disorders." ;
RN Cell 85:597-605(1996) .
RN [9]
RN REVIEW ON MFS VARIANTS.
RP MEDLINE=96174615; PubMed=8594563;
RA Colloid G., Beroud C., Soussi T., Junien C., Boileau C.;
RT "Software and database for the analysis of mutations in the human
RT FBN1 gene." ;
RN Nucleic Acids Res. 24:137-141(1996) .
RN [10]
RN REVIEW ON MFS VARIANTS.
RP MEDLINE=97169383; PubMed=9016526;
RA Colloid-Beroud G., Beroud C., Ales L., Black C., Boxer M., Brock D.J.,
RA Godfrey M., Hayward C., Karttunen L., Milewicz D., Peltonen L.,
RA Richards R.I., Wang W., Junten C., Boileau C.;
RT "Marfan Database (second edition): software and database for the
RT analysis of mutations in the human FBN1 gene." ;
RN Nucleic Acids Res. 25:147-150(1997) .
RN [11]
RN REVIEW ON VARIANTS.
RP MEDLINE=98062175; PubMed=9401003;
RA Hayward C., Brock D.J.H.;
RT "Fibrillin-1 mutations in Marfan syndrome and other type-1
RT fibrillogenopathies." ;
RN Hum. Mutat. 10:415-423(1997) .
RN [12]
RN VARIANT MFS PRO-1137.
RP MEDLINE=91304569; PubMed=1852208;
RA Dietz H.C., Cutting G.R., Pyeritz R.E., Maslen C.L., Sakai L.Y.,
RA Corson G.M., Puffenberger E.G., Hamosh A., Nanthakumar E.J.,
RA Curristin S.M., Stetten G., Meyers D.A., Francomano C.A.;
RT "Marfan syndrome caused by a recurrent de novo missense mutation in
RT the fibrillin gene." ;
RN Nature 352:337-339(1991) .
RN [13]
RN VARIANTS MFS SER-1249; ANG²1663; SER-2221 AND SER-2307.
RP MEDLINE=93250834; PubMed=1301946;
RA Dietz H.C., Sariva J.M., Pyeritz R.E., Cutting G.R., Francomano C.A.,
RT "Clustering of fibrillin (FBN1) missense mutations in Marfan syndrome
RT patients at cysteine residues in EGF-like domains." ;
RN Hum. Mutat. 1:366-374(1992) .
RN [14]
RN VARIANT MFS SER-2307.
RP MEDLINE=92235290; PubMed=1569206;
RA Dietz H.C., Pyeritz R.E., Puffenberger E.G., Kendzior R.J. Jr.,
RA Corson G.M., Maslen C.L., Sakai L.Y., Francomano C.A., Cutting G.R.;
RT "Marfan phenotype variability in a family segregating a missense
RT mutation in the epidermal growth factor-like motif of the fibrillin
RT gene." ;
RN J. Clin. Invest. 89:1674-1680(1992) .
RN [15]
RN VARIANTS MFS ILE-548 AND ALA-723.
RP MEDLINE=94010946; PubMed=13046497;
RA Dietz H.C., McIntosh I., Sakai L.Y., Corson G.M., Chalberg S.C.,
RA Pyeritz R.E., Francomano C.A.;
RT "Four novel FBN1 mutations: significance for mutant transcript level
RT and EGF-like domain calcium binding in the pathogenesis of Marfan
RT syndrome." ;
RN Genomics 17:468-475(1993) .
RN [16]
RN VARIANTS MFS SER-2144.
RP MEDLINE=93278402; PubMed=8504310;
RA Hewett D.R., Lynch J.R., Smith R., Sykes B.C.;
RT "A novel fibrillin mutation in the Marfan syndrome which could
RT disrupt calcium binding of the epidermal growth factor-like module." ;
RN Hum. Mol. Genet. 2:475-477(1993) .
RN [17]
RN VARIANTS MFS R-862; Y-1137; P-1137 AND F-1589, AND VARIANT A-1148.
RP MEDLINE=94108431; PubMed=8281141;
RA Tynan K., Comeau K., Pearson M., Wilgenbus P., Levitt D., Gasner C.,
RA Berg M.A., Miller D.C., Francke U.;

RT "Mutation screening of complete fibrillin-1 coding sequence: report
RT of five new mutations, including two in 8-cysteine domains".;
RN Hum. Mol. Genet. 2:1813-1821(1993).
RN [18]
RP VARIANTS MFS GLY-217 AMD ARG-2627.
RX MEDLINE=95067970; PubMed=7977366;
RA Karttunen L., Raghunath M., Loengqvist L., Peltonen L.;
RT "A compound heterozygous Marfan patient: two defective fibrillin
RL alleles result in a lethal phenotype.";
RN Am. J. Hum. Genet. 55:1083-1091(1994).
RN [19]
RP VARIANT EL IYS-2447.
RX MEDLINE=94245249; PubMed=8188302;
RA Lomqvist L., Child A., Kainulainen K., Davidson R., Puhakka L.,
RA Peltonen L.;
RT "A novel mutation of the fibrillin gene causing ectopia lentis".;
RL Genomics 19:573-576(1994).
RN [20]
RP VARIANT MFS CYS-627.
RX MEDLINE=94272487; PubMed=8004112;
RA Hayward C., Rae A.L., Porteous M.E.M., Logie L.J., Brock L.J.;
RT "Two novel mutations and a neutral polymorphism in EGF-like domains
of the fibrillin gene (FBN1): SSCP screening of exons 15-21 in Marfan
RT syndrome patients".;
RL Hum. Mol. Genet. 3:373-375(1994).
RN [21]
RP VARIANT MFS CYS-122.
RX MEDLINE=94314577; PubMed=8040326;
RA Stahl-Hallengren C., Ukkonen T., Kainulainen K., Kristoffersson U.,
RA Saxne T., Torqvist K., Peltonen L.;
RT "An extra cysteine in one of the non-calcium-binding epidermal growth
RT factor-like motifs of the FBN1 polypeptide is connected to a novel
RL variant of Marfan syndrome".;
RN J. Clin. Invest. 94:709-713(1994).
RN [22]
RP VARIANT MFS TYR-1223.
RX MEDLINE=94351682; PubMed=8071963;
RA Hewett D.R., Lynch J.R., Child A., Sykes B.C.;
RT "A new missense mutation of fibrillin in a patient with Marfan
RT syndrome".;
RN J. Med. Genet. 31:338-339(1994).
RN [23]
RP VARIANT MFS HIS-1170.
RX MEDLINE=95174777; PubMed=7870075;
RA Hayward C., Porteous M.E.M., Brock D.J.H.;
RT "A novel mutation in the fibrillin gene (FBN1) in familial
RL aachondrocyly".;
RN Mol. Cell. Probes 8:325-327(1994).
RN [24]
RP VARIANTS MFS G217;N1023;A1074;Y1242;R1513;E2127;W2151;K2447 AND R2511.
RX MEDLINE=94184368; PubMed=8136837;
RA Kainulainen K., Karttunen L., Puhakka L., Sakai L., Peltonen L.;
RT "Mutations in the fibrillin gene responsible for dominant ectopia
RT lentis and neonatal Marfan syndrome".;
RN Nat. Genet. 6:64-69(1994).
RN [25]
RP VARIANT SER-1127.
RX MEDLINE=95282774; PubMed=7762551;

| | | | | | |
|----|--|--|------------------|-------------|--------------|
| | Query Match | 12.4% | Score 363; | DB 1; | Length 2871; |
| | Best Local Similarity | 30.8%; | Pred. No. 2e-16; | | |
| | Matches 111; Conservative | 22; | Mismatches 105; | Indels 122; | Gaps 20; |
| QY | 244 DCSVNGCCGCAACNAIPAPRCQCAGALQADGSCTASATQSCNDLCEHFCVPNP--- | 300 | | | |
| | :::: : | : | ::: | : | |
| Db | 1200 ECSIMGCGTEFCTSESSYECSQPGFALMPDRKST-----DIDE--CEDNPNIC | 1249 | | | |
| QY | 301 -----DDPGSYSCSGCEGYRLAAQHHCEDVDCLLESPPC-PQRCVNTQGFGFCHCP | 353 | | | |
| | Db | 1250 DGGOCTNTPGEYRCLCYGGFMASEDKTCVYNEDLNPNICLSGTCENTKGSLTCHCDM | 1309 | | |
| QY | 354 NYDLVDGE--CYEPDPDC--FRANEYQQQLNQI--SYLCVCABEFA-----PIP | 398 | | | |
| | : : : : | | : : : : | | |

```

Db 1310 GYGGKRGKTCCTD-INCEIGAHNGKHAVCNTAGSFKSCSPGWIGDGIKTDLDECS 1368
QY 399 HEPHRCOMFCNQTACPADCDPNTQAS--CCPCRGYIILDDEFCIDDEC-EN----- 447
Db 1369 NGHMCSSQH-----ADC-KNTMGSTRCLCKESY-TGDEPTCTDDECEENLWLCNGQ 1419
QY 448 ----GGF-----CS-----GVCHNLPTGFCICGPDASLV 473
Db 1420 CLNAPGGYRCEDMGFVPASADKACEDIDRCSLPNICVFPTCHNLPOLFRCCEIGELD 1479
QY 474 RHIG-----IDCSGKXVDGDSGSGEPPSPPTGS-----TLAPPAGLV 513
Db 1480 RSGGNCCTVNECLDPTTCISGNCVN-----TPGSYICDPPDELPNTRVGCV 1527

RESULT 10
FBN1_MOUSE STANDARD; PRT; 2871 AA.
ID FBN1_MOUSE
AC 061554; 060826;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FIBRILIN 1 PRECURSOR.
GN FBN1 OR FBN-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=95130561; PubMed=7829516;
RA Yin W., Gemmiller J., Sanguinetti C., Smiley E., Pangillan T.,
RA Pereira L., Ramirez F., Bonadio J.;
RT "Primary structure and developmental expression of Fbn-1, the mouse
RT fibrillin gene."
RL J. Biol. Chem. 270:1798-1806(1995).
RN 12
RP SEQUENCE FROM N.A.
RA STRAIN=CD-1; TISSUE=Kidney;
RA Ota K., Kumar A., Wada J., Liu Z., Kanwar Y.S.;
RL Submitted (Apr-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILIN-1-CONTAINING MICROFIBRILS PROVIDE
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -1- FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC FIBRILIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC MICROFIBRILS (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L29454; AAA56840.1; -
DR EMBL; U22493; AAA64217.1; -
DR HSSP; P35555; IAPJ.
DR MGD; MGI:95489; Fbn1.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR002212; TB.
DR Pfam; PF00008; EGF; 46.
DR Pfam; PF00683; TB; 9.
DR SMART; SM00179; EGF_Ca; 42.
DR SMART; SM00001; EGF_like; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; 2.

```

```

DR PROSITE; PS01186; EGF_2; 38.
DR PROSITE; PS01187; EGF_CA; 45.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1
FT CHAIN 28
FT DOMAIN 81
FT DOMAIN 112
FT DOMAIN 115
FT DOMAIN 147
FT DOMAIN 147
FT DOMAIN 246
FT DOMAIN 288
FT DOMAIN 329
FT DOMAIN 330
FT DOMAIN 401
FT DOMAIN 402
FT DOMAIN 449
FT DOMAIN 449
FT DOMAIN 490
FT DOMAIN 529
FT DOMAIN 530
FT DOMAIN 571
FT DOMAIN 572
FT DOMAIN 612
FT DOMAIN 612
FT DOMAIN 653
FT DOMAIN 656
FT DOMAIN 723
FT DOMAIN 764
FT DOMAIN 765
FT DOMAIN 807
FT DOMAIN 810
FT DOMAIN 910
FT DOMAIN 952
FT DOMAIN 1018
FT DOMAIN 1028
FT DOMAIN 1070
FT DOMAIN 1112
FT DOMAIN 1113
FT DOMAIN 1155
FT DOMAIN 1197
FT DOMAIN 1238
FT DOMAIN 1279
FT DOMAIN 1280
FT DOMAIN 1321
FT DOMAIN 1362
FT DOMAIN 1363
FT DOMAIN 1404
FT DOMAIN 1445
FT DOMAIN 1446
FT DOMAIN 1487
FT DOMAIN 1527
FT DOMAIN 1558
FT DOMAIN 1599
FT DOMAIN 1647
FT DOMAIN 1668
FT DOMAIN 1688
FT DOMAIN 1699
FT DOMAIN 1766
FT DOMAIN 1807
FT DOMAIN 1848
FT DOMAIN 1849
FT DOMAIN 1890
FT DOMAIN 1891
FT DOMAIN 1929
FT DOMAIN 1930
FT DOMAIN 1973
FT DOMAIN 2012
FT DOMAIN 2013
FT DOMAIN 2054
FT DOMAIN 2055
FT DOMAIN 2121
FT DOMAIN 2137
FT DOMAIN 2166
FT DOMAIN 2205
FT DOMAIN 2206
FT DOMAIN 2247
FT DOMAIN 2290
FT DOMAIN 2291
FT DOMAIN 2332
FT DOMAIN 2333
FT DOMAIN 2402
FT DOMAIN 2443
FT DOMAIN 2444
FT DOMAIN 2484
FT DOMAIN 2523
FT DOMAIN 2524
FT DOMAIN 2566
FT DOMAIN 2567
FT DOMAIN 2606
FT DOMAIN 2607
FT DOMAIN 2647
FT DOMAIN 2648
FT DISULFID 85
FT DISULFID 89
FT DISULFID 94
FT DISULFID 100
FT DISULFID 102
FT DISULFID 111
FT DISULFID 119
FT DISULFID 123
FT DISULFID 134
FT DISULFID 145
FT DISULFID 150
FT DISULFID 160
FT DISULFID 166
FT DISULFID 177
FT DISULFID 188
FT DISULFID 250
FT DISULFID 262
FT DISULFID 271
FT DISULFID 273
FT DISULFID 286
FT DISULFID 286

```

| | QY | 212 | SAANVAFGL-----QLMKTAPPGAVGSHAKAREAPGAMDSYENGSCCHACNAINPAPPC | 265 |
|--------|---|---|--|---------|
| | | 1162 | SAALCHGRCVNLIGKYOCACNPNGHPTNDRFLFCVDIDECSDIMNGCFETCTNSDGSYEC | 1221 |
| | QY | 266 | QCPAGAAIADNRSGTASATGSCNDLCEHFVMP-----DQGSYSCCEGTGRU | 316 |
| | Db | 1222 | SCQPGALMPDORST-----DIDQ--CEDNPNICDGGQCTNPITGFEKCLCTGYFMA | 1271 |
| | QY | 317 | AADHRCEDVDLCLEPSC-PQRCVNTQGFEGCHYPNYDLVDE--CVEPVPDC--FR | 371 |
| | Db | 1272 | SEDMKTCVNVNECDLNPNICISGCTEENTKSGFICHCDMGYSGKKGKGTCTD-INCEIGA | 1330 |
| | QY | 372 | ANCEYOCQPLNQT-SYLCACAGFA-----PIPHPRHCQMCNQTACPAQCDPN | 420 |
| | Db | 1331 | HNGGRHACVTNTPAGSKCSPGMIGDDIKCTDLDECNSGTHMCSQH-----ADC-KN | 1382 |
| | QY | 421 | TOAS--CEGPEGYLLDGFICTDIDEC-EN-----GGF----- | 450 |
| | Db | 1383 | TWGSRYCLCKDGY-FTDGFPTCTDLDECSENILNCGNGCCLNAPGGRHCECMGFVPSADG | 1441 |
| | QY | 451 | -----CS-----GCHNLPRGTFFECICGPPSALYRHIG-----TTDDSKK | 484 |
| | Db | 1442 | KACEDIDESLPIVCVFGCHNLPGLFKECEIGEYEDRSGGCTDYNECLDPTTCISGN | 1501 |
| | QY | 485 | VDGDSGSGEPSPPPGCS-----TTPPAVGLY | 513 |
| | Db | 1502 | CVN-----TPGSYTCDCSPDELPNTRGCV | 1527 |
| RESULT | 11 | | | |
| MTN2 | HUMAN | STANDARD: | PRT: | 956 AA. |
| AC | 000339: | Q9NSZ1: | | |
| DT | 30-MAY-2000 | (Rel. 39, Created) | | |
| DT | 20-AUG-2001 | (Rel. 40, Last sequence update) | | |
| DT | 20-AUG-2001 | (Rel. 40, Last annotation update) | | |
| DE | MATRILIN-2 | PRECURSOR. | | |
| GN | MATN2. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RP | RP | SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS). | | |
| RP | RP | PubMed-11124542; | | |
| RA | PubMed-11124542; | | | |
| RA | Kiss I., Deak F.; | | | |
| RT | "Primary structure of human matrilin-2, chromosome location of the | | | |
| RT | MATN2 gene and conservation of an A+T-rich intron in matrilin genes." | | | |
| RL | Cytogenet. Cell Genet. 90:323-327(2000). | | | |
| RN | 121 | | | |
| RP | SEQUENCE OF 644-956 FROM N.A. | | | |
| RP | MEDLINE=97238863: PubMed=9083061: | | | |
| RA | Deak F., Plecha D., Bachratl C., Paulsson M., Kiss I.; | | | |
| RT | "Primary structure and expression of matrilin-2, the closest relative | | | |
| RT | of cartilage matrix protein within the von Willebrand factor type A- | | | |
| RL | like module superfamily." | | | |
| RL | J. Biol. Chem. 272:9268-9274(1997). | | | |
| RN | 131 | | | |
| RP | SEQUENCE OF 244-956 FROM N.A. | | | |
| RC | TISUD-testis; | | | |
| RC | Duesterhoelt A., Lauber J., Mewes H.-W., Gassenhuber J., Wiemann S.; | | | |
| RL | Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases. | | | |
| CC | -1- FUNCTION: INVOLVED IN MATRIX ASSEMBLY (BY SIMILARITY). | | | |
| CC | -1- SUBCELLULAR LOCATION: SECRETED. | | | |
| CC | -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A | | | |
| CC | SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING. | | | |
| CC | -1- SIMILARITY: CONTAINS 10 EGF-LIKE DOMAINS. | | | |

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U69263; AAC51260.2; .
EMBL: AL137638; CAB70853.1; ALT_INIT.
HSSP: P05099; IAO5.
MIM: 602108; .
InterPro: IPR000152; Asx_hydroxyl.
InterPro: IPR000561; EGF-like.
InterPro: IPR001881; EGF_Ca.
InterPro: IPR002035; WMFA.
Pfam: PF00008; EGF; 10.
Pfam: PF00092; vwa; 1.
PRINTS: PR00453; VWFADOMAIN.
SMART: SM00179; EGF_Ca; 1.
SMART: SM00001; EGF_Like; 9.
PROSITE: PS00010; ASX_HYDROXYL; 9.
PROSITE: PS01186; EGF_2; 9.
PROSITE: PS50234; WMFA; 2.
EGF-like domain; signal; Glycoprotein; Repeat; coiled coil; Alternative splicing.
SIGNA 1
CHAIN 24 956
DOMAIN 57 232
DOMAIN 238 278
DOMAIN 279 319
DOMAIN 320 360
DOMAIN 361 401
DOMAIN 402 442
DOMAIN 443 483
DOMAIN 484 524
DOMAIN 525 565
DOMAIN 566 606
DOMAIN 607 647
DOMAIN 648 688
DOMAIN 689 729
DOMAIN 730 770
DOMAIN 771 811
DOMAIN 812 852
DOMAIN 853 893
DOMAIN 894 934
DOMAIN 935 975
DOMAIN 976 1016
DOMAIN 1017 1057
DOMAIN 1058 1098
DOMAIN 1099 1139
DOMAIN 1140 1180
DOMAIN 1181 1221
DOMAIN 1222 1262
DOMAIN 1263 1303
DOMAIN 1304 1344
DOMAIN 1345 1385
DOMAIN 1386 1426
DOMAIN 1427 1467
DOMAIN 1468 1508
DOMAIN 1509 1549
DOMAIN 1550 1590
DOMAIN 1591 1631
DOMAIN 1632 1672
DOMAIN 1673 1713
DOMAIN 1714 1754
DOMAIN 1755 1795
DOMAIN 1796 1836
DOMAIN 1837 1877
DOMAIN 1878 1918
DOMAIN 1919 1959
DOMAIN 1960 2000
DOMAIN 2001 2041
DOMAIN 2042 2082
DOMAIN 2083 2123
DOMAIN 2124 2164
DOMAIN 2165 2205
DOMAIN 2206 2246
DOMAIN 2247 2287
DOMAIN 2288 2328
DOMAIN 2329 2369
DOMAIN 2370 2410
DOMAIN 2411 2451
DOMAIN 2452 2492
DOMAIN 2493 2533
DOMAIN 2534 2574
DOMAIN 2575 2615
DOMAIN 2616 2656
DOMAIN 2657 2697
DOMAIN 2698 2738
DOMAIN 2739 2779
DOMAIN 2780 2820
DOMAIN 2821 2861
DOMAIN 2862 2902
DOMAIN 2903 2943
DOMAIN 2944 2984
DOMAIN 2985 3025
DOMAIN 3026 3066
DOMAIN 3067 3107
DOMAIN 3108 3148
DOMAIN 3149 3189
DOMAIN 3190 3230
DOMAIN 3231 3271
DOMAIN 3272 3312
DOMAIN 3313 3353
DOMAIN 3354 3394
DOMAIN 3395 3435
DOMAIN 3436 3476
DOMAIN 3477 3517
DOMAIN 3518 3558
DOMAIN 3559 3599
DOMAIN 3600 3640
DOMAIN 3641 3681
DOMAIN 3682 3722
DOMAIN 3723 3763
DOMAIN 3764 3804
DOMAIN 3805 3845
DOMAIN 3846 3886
DOMAIN 3887 3927
DOMAIN 3928 3968
DOMAIN 3969 4009
DOMAIN 4010 4050
DOMAIN 4051 4091
DOMAIN 4092 4132
DOMAIN 4133 4173
DOMAIN 4174 4214
DOMAIN 4215 4255
DOMAIN 4256 4296
DOMAIN 4297 4337
DOMAIN 4338 4378
DOMAIN 4379 4419
DOMAIN 4420 4460
DOMAIN 4461 4501
DOMAIN 4502 4542
DOMAIN 4543 4583
DOMAIN 4584 4624
DOMAIN 4625 4665
DOMAIN 4666 4706
DOMAIN 4707 4747
DOMAIN 4748 4788
DOMAIN 4789 4829
DOMAIN 4830 4870
DOMAIN 4871 4911
DOMAIN 4912 4952
DOMAIN 4953 4993
DOMAIN 4994 5034
DOMAIN 5035 5075
DOMAIN 5076 5116
DOMAIN 5117 5157
DOMAIN 5158 5198
DOMAIN 5199 5239
DOMAIN 5240 5280
DOMAIN 5281 5321
DOMAIN 5322 5362
DOMAIN 5363 5403
DOMAIN 5404 5444
DOMAIN 5445 5485
DOMAIN 5486 5526
DOMAIN 5527 5567
DOMAIN 5568 5608
DOMAIN 5609 5649
DOMAIN 5650 5690
DOMAIN 5691 5731
DOMAIN 5732 5772
DOMAIN 5773 5813
DOMAIN 5814 5854
DOMAIN 5855 5895
DOMAIN 5896 5936
DOMAIN 5937 5977
DOMAIN 5978 6018
DOMAIN 6019 6059
DOMAIN 6060 6099
DOMAIN 6100 6140
DOMAIN 6141 6181
DOMAIN 6182 6222
DOMAIN 6223 6263
DOMAIN 6264 6304
DOMAIN 6305 6345
DOMAIN 6346 6386
DOMAIN 6387 6427
DOMAIN 6428 6468
DOMAIN 6469 6509
DOMAIN 6510 6550
DOMAIN 6551 6591
DOMAIN 6592 6632
DOMAIN 6633 6673
DOMAIN 6674 6714
DOMAIN 6715 6755
DOMAIN 6756 6796
DOMAIN 6797 6837
DOMAIN 6838 6878
DOMAIN 6879 6919
DOMAIN 6920 6960
DOMAIN 6961 7001
DOMAIN 7002 7042
DOMAIN 7043 7083
DOMAIN 7084 7124
DOMAIN 7125 7165
DOMAIN 7166 7206
DOMAIN 7207 7247
DOMAIN 7248 7288
DOMAIN 7289 7329
DOMAIN 7330 7370
DOMAIN 7371 7411
DOMAIN 7412 7452
DOMAIN 7453 7493
DOMAIN 7494 7534
DOMAIN 7535 7575
DOMAIN 7576 7616
DOMAIN 7617 7657
DOMAIN 7658 7698
DOMAIN 7699 7739
DOMAIN 7740 7780
DOMAIN 7781 7821
DOMAIN 7822 7862
DOMAIN 7863 7903
DOMAIN 7904 7944
DOMAIN 7945 7985
DOMAIN 7986 8026
DOMAIN 8027 8067
DOMAIN 8068 8108
DOMAIN 8109 8149
DOMAIN 8150 8190
DOMAIN 8191 8231
DOMAIN 8232 8272
DOMAIN 8273 8313
DOMAIN 8314 8354
DOMAIN 8355 8395
DOMAIN 8396 8436
DOMAIN 8437 8477
DOMAIN 8478 8518
DOMAIN 8519 8559
DOMAIN 8560 8600
DOMAIN 8601 8641
DOMAIN 8642 8682
DOMAIN 8683 8723
DOMAIN 8724 8764
DOMAIN 8765 8805
DOMAIN 8806 8846
DOMAIN 8847 8887
DOMAIN 8888 8928
DOMAIN 8929 8969
DOMAIN 8970 9010
DOMAIN 9011 9051
DOMAIN 9052 9092
DOMAIN 9093 9133
DOMAIN 9134 9174
DOMAIN 9175 9215
DOMAIN 9216 9256
DOMAIN 9257 9297
DOMAIN 9298 9338
DOMAIN 9339 9379
DOMAIN 9380 9420
DOMAIN 9421 9461
DOMAIN 9462 9502
DOMAIN 9503 9543
DOMAIN 9544 9584
DOMAIN 9585 9625
DOMAIN 9626 9666
DOMAIN 9667 9707
DOMAIN 9708 9748
DOMAIN 9749 9789
DOMAIN 9790 9830
DOMAIN 9831 9871
DOMAIN 9872 9912
DOMAIN 9913 9953
DOMAIN 9954 9994
DOMAIN 10000 10000
DOMAIN 10001 10001
DOMAIN 10002 10002
DOMAIN 10003 10003
DOMAIN 10004 10004
DOMAIN 10005 10005
DOMAIN 10006 10006
DOMAIN 10007 10007
DOMAIN 10008 10008
DOMAIN 10009 10009
DOMAIN 10010 10010
DOMAIN 10011 10011
DOMAIN 10012 10012
DOMAIN 10013 10013
DOMAIN 10014 10014
DOMAIN 10015 10015
DOMAIN 10016 10016
DOMAIN 10017 10017
DOMAIN 10018 10018
DOMAIN 10019 10019
DOMAIN 10020 10020
DOMAIN 10021 10021
DOMAIN 10022 10022
DOMAIN 10023 10023
DOMAIN 10024 10024
DOMAIN 10025 10025
DOMAIN 10026 10026
DOMAIN 10027 10027
DOMAIN 10028 10028
DOMAIN 10029 10029
DOMAIN 10030 10030
DOMAIN 10031 10031
DOMAIN 10032 10032
DOMAIN 10033 10033
DOMAIN 10034 10034
DOMAIN 10035 10035
DOMAIN 10036 10036
DOMAIN 10037 10037
DOMAIN 10038 10038
DOMAIN 10039 10039
DOMAIN 10040 10040
DOMAIN 10041 10041
DOMAIN 10042 10042
DOMAIN 10043 10043
DOMAIN 10044 10044
DOMAIN 10045 10045
DOMAIN 10046 10046
DOMAIN 10047 10047
DOMAIN 10048 10048
DOMAIN 10049 10049
DOMAIN 10050 10050
DOMAIN 10051 10051
DOMAIN 10052 10052
DOMAIN 10053 10053
DOMAIN 10054 10054
DOMAIN 10055 10055
DOMAIN 10056 10056
DOMAIN 10057 10057
DOMAIN 10058 10058
DOMAIN 10059 10059
DOMAIN 10060 10060
DOMAIN 10061 10061
DOMAIN 10062 10062
DOMAIN 10063 10063
DOMAIN 10064 10064
DOMAIN 10065 10065
DOMAIN 10066 10066
DOMAIN 10067 10067
DOMAIN 10068 10068
DOMAIN 10069 10069
DOMAIN 10070 10070
DOMAIN 10071 10071
DOMAIN 10072 10072
DOMAIN 10073 10073
DOMAIN 10074 10074
DOMAIN 10075 10075
DOMAIN 10076 10076
DOMAIN 10077 10077
DOMAIN 10078 10078
DOMAIN 10079 10079
DOMAIN 10080 10080
DOMAIN 10081 10081
DOMAIN 10082 10082
DOMAIN 10083 10083
DOMAIN 10084 10084
DOMAIN 10085 10085
DOMAIN 10086 10086
DOMAIN 10087 10087
DOMAIN 10088 10088
DOMAIN 10089 10089
DOMAIN 10090 10090
DOMAIN 10091 10091
DOMAIN 10092 10092
DOMAIN 10093 10093
DOMAIN 10094 10094
DOMAIN 10095 10095
DOMAIN 10096 10096
DOMAIN 10097 10097
DOMAIN 10098 10098
DOMAIN 10099 10099
DOMAIN 10100 10100
DOMAIN 10101 10101
DOMAIN 10102 10102
DOMAIN 10103 10103
DOMAIN 10104 10104
DOMAIN 10105 10105
DOMAIN 10106 10106
DOMAIN 10107 10107
DOMAIN 10108 10108
DOMAIN 10109 10109
DOMAIN 10110 10110
DOMAIN 10111 10111
DOMAIN 10112 10112
DOMAIN 10113 10113
DOMAIN 10114 10114
DOMAIN 10115 10115
DOMAIN 10116 10116
DOMAIN 10117 10117
DOMAIN 10118 10118
DOMAIN 10119 10119
DOMAIN 10120 10120
DOMAIN 10121 10121
DOMAIN 10122 10122
DOMAIN 10123 10123
DOMAIN 10124 10124
DOMAIN 10125 10125
DOMAIN 10126 10126
DOMAIN 10127 10127
DOMAIN 10128 10128
DOMAIN 10129 10129
DOMAIN 10130 10130
DOMAIN 10131 10131
DOMAIN 10132 10132
DOMAIN 10133 10133
DOMAIN 10134 10134
DOMAIN 10135 10135
DOMAIN 10136 10136
DOMAIN 10137 10137
DOMAIN 10138 10138
DOMAIN 10139 10139
DOMAIN 10140 10140
DOMAIN 10141 10141
DOMAIN 10142 10142
DOMAIN 10143 10143
DOMAIN 10144 10144
DOMAIN 10145 10145
DOMAIN 10146 10146
DOMAIN 10147 10147
DOMAIN 10148 10148
DOMAIN 10149 10149
DOMAIN 10150 10150
DOMAIN 10151 10151
DOMAIN 10152 10152
DOMAIN 10153 10153
DOMAIN 10154 10154
DOMAIN 10155 10155
DOMAIN 10156 10156
DOMAIN 10157 10157
DOMAIN 10158 10158
DOMAIN 10159 10159
DOMAIN 10160 10160
DOMAIN 10161 10161
DOMAIN 10162 10162
DOMAIN 10163 10163
DOMAIN 10164 10164
DOMAIN 10165 10165
DOMAIN 10166 10166
DOMAIN 10167 10167
DOMAIN 10168 10168
DOMAIN 10169 10169
DOMAIN 10170 10170
DOMAIN 10171 10171
DOMAIN 10172 10172
DOMAIN 10173 10173
DOMAIN 10174 10174
DOMAIN 10175 10175
DOMAIN 10176 10176
DOMAIN 10177 10177
DOMAIN 10178 10178
DOMAIN 10179 10179
DOMAIN 10180 10180
DOMAIN 10181 10181
DOMAIN 10182 10182
DOMAIN 10183 10183
DOMAIN 10184 10184
DOMAIN 10185 10185
DOMAIN 10186 10186
DOMAIN 10187 10187
DOMAIN 10188 10188
DOMAIN 10189 10189
DOMAIN 10190 10190
DOMAIN 10191 10191
DOMAIN 10192 10192
DOMAIN 10193 10193
DOMAIN 10194 10194
DOMAIN 10195 10195
DOMAIN 10196 10196
DOMAIN 10197 10197
DOMAIN 10198 10198
DOMAIN 10199 10199
DOMAIN 10200 10200
DOMAIN 10201 10201
DOMAIN 10202 10202
DOMAIN 10203 10203
DOMAIN 10204 10204
DOMAIN 10205 10205
DOMAIN 10206 10206
DOMAIN 10207 10207
DOMAIN 10208 10208
DOMAIN 10209 10209
DOMAIN 10210 10210
DOMAIN 10211 10211
DOMAIN 10212 10212
DOMAIN 10213 10213
DOMAIN 10214 10214
DOMAIN 10215 10215
DOMAIN 10216 10216
DOMAIN 10217 10217
DOMAIN 10218 10218
DOMAIN 10219 10219
DOMAIN 10220 10220
DOMAIN 10221 10221
DOMAIN 10222 10222
DOMAIN 10223 10223
DOMAIN 10224 10224
DOMAIN 10225 10225
DOMAIN 10226 10226
DOMAIN 10227 10227
DOMAIN 10228 10228
DOMAIN 10229 10229
DOMAIN 10230 10230
DOMAIN 10231 10231
DOMAIN 10232 10232
DOMAIN 10233 10233
DOMAIN 10234 10234
DOMAIN 10235 10235
DOMAIN 10236 10236
DOMAIN 10237 10237
DOMAIN 10238 10238
DOMAIN 10239 10239
DOMAIN 10240 10240
DOMAIN 10241 10241
DOMAIN 10242 10242
DOMAIN 10243 10243
DOMAIN 10244 10244
DOMAIN 10245 10245
DOMAIN 10246 10246
DOMAIN 10247 10247
DOMAIN 10248 10248
DOMAIN 10249 10249
DOMAIN 10250 10250
DOMAIN 10251 10251
DOMAIN 10252 10252
DOMAIN 10253 10253
DOMAIN 10254 10254
DOMAIN 10255 10255
DOMAIN 10256 10256
DOMAIN 10257 10257
DOMAIN 10258 10258
DOMAIN 10259 10259
DOMAIN 10260 10260
DOMAIN 10261 10261
DOMAIN 10262 10262
DOMAIN 10263 10263
DOMAIN 10264 10264
DOMAIN 10265 10265
DOMAIN 10266 10266
DOMAIN 10267 10267
DOMAIN 10268 10268
DOMAIN 10269 10269
DOMAIN 10270 10270
DOMAIN 10271 10271
DOMAIN 10272 10272
DOMAIN 10273 10273
DOMAIN 10274 10274
DOMAIN 10275 10275
DOMAIN 10276 10276
DOMAIN 10277 10277
DOMAIN 10278 10278
DOMAIN 10279 10279
DOMAIN 10280 10280
DOMAIN 10281 10281
DOMAIN 10282 10282
DOMAIN 10283 10283
DOMAIN 10284 10284
DOMAIN 10285 10285
DOMAIN 10286 10286
DOMAIN 10287 10287
DOMAIN 10288 10288
DOMAIN 10289 10289
DOMAIN 10290 10290
DOMAIN 10291 10291
DOMAIN 10292 10292
DOMAIN 10293 10293
DOMAIN 10294 10294
DOMAIN 10295 10295
DOMAIN 10296 10296
DOMAIN 10297 10297
DOMAIN 10298 10298
DOMAIN 10299 10299
DOMAIN 10300 10300
DOMAIN 10301 10301
DOMAIN 10302 10302
DOMAIN 10303 10303
DOMAIN 10304 10304
DOMAIN 10305 10305
DOMAIN 10306 10306
DOMAIN 10307 10307
DOMAIN 10308 10308
DOMAIN 10309 10309
DOMAIN 10310 10310
DOMAIN 10311 10311
DOMAIN 10312 10312
DOMAIN 10313 10313
DOMAIN 10314 10314
DOMAIN 10315 10315
DOMAIN 10316 10316
DOMAIN 10317 10317
DOMAIN 10318 10318
DOMAIN 10319 10319
DOMAIN 10320 10320
DOMAIN 10321 10321
DOMAIN 10322 10322
DOMAIN 10323 10323
DOMAIN 10324 10324
DOMAIN 10325 10325
DOMAIN 10326 10326
DOMAIN 10327 10327
DOMAIN 10328 10328
DOMAIN 10329 10329
DOMAIN 10330 10330
DOMAIN 10331 10331
DOMAIN 10332 10332
DOMAIN 10333 10333
DOMAIN 10334 10334
DOMAIN 10335 10335
DOMAIN 10336 10336
DOMAIN 10337 10337
DOMAIN 10338 10338
DOMAIN 10339 10339
DOMAIN 10340 10340
DOMAIN 10341 10341
DOMAIN 10342 10342
DOMAIN 10343 10343
DOMAIN 10344 10344
DOMAIN 10345 10345
DOMAIN 10346 10346
DOMAIN 10347 10347
DOMAIN 10348 10348
DOMAIN 10349 10349
DOMAIN 10350 10350
DOMAIN 10351 10351
DOMAIN 10352 10352
DOMAIN 10353 10353
DOMAIN 10354 10354
DOMAIN 10355 10355
DOMAIN 10356 10356
DOMAIN 10357 10357
DOMAIN 10358 10358
DOMAIN 10359 10359
DOMAIN 10360 10360
DOMAIN 10361 10361
DOMAIN 10362 10362
DOMAIN 10363 10363
DOMAIN 10364 10364
DOMAIN 10365 10365
DOMAIN 10366 10366
DOMAIN 10367 10367
DOMAIN 10368 10368
DOMAIN 10369 10369
DOMAIN 10370 10370
DOMAIN 10371 10371
DOMAIN 10372 10372
DOMAIN 10373 10373
DOMAIN 10374 10374
DOMAIN 10375 10375
DOMAIN 10376 10376
DOMAIN 10377 10377
DOMAIN 10378 10378
DOMAIN 10379 10379
DOMAIN 10380 10380
DOMAIN 10381 10381
DOMAIN 10382 10382
DOMAIN 10383 10383
DOMAIN 10384 10384
DOMAIN 10385 10385
DOMAIN 10386 10386
DOMAIN 10387 10387
DOMAIN 10388 10388
DOMAIN 10389 10389
DOMAIN 10390 10390
DOMAIN 10391 10391
DOMAIN 10392 10392
DOMAIN 10393 10393
DOMAIN 10394 10394
DOMAIN 10395 10395
DOMAIN 10396 10396
DOMAIN 10397 10397
DOMAIN 10398 10398
DOMAIN 10399 10399
DOMAIN 10400 10400
DOMAIN 10401 10401
DOMAIN 10402 10402
DOMAIN 10403 10403
DOMAIN 10404 10404
DOMAIN 10405 10405
DOMAIN 10406 10406
DOMAIN 10407 10407
DOMAIN 10408 10408
DOMAIN 10409 10409
DOMAIN 10410 10410
DOMAIN 10411 10411
DOMAIN 10412 10412
DOMAIN 10413 10413
DOMAIN 10414 10414
DOMAIN 10415 10415
DOMAIN 10416 10416
DOMAIN 10417 10417
DOMAIN 10418 10418
DOMAIN 10419 10419
DOMAIN 10420 10420
DOMAIN 10421 10421
DOMAIN 10422 10422
DOMAIN 10423 10423
DOMAIN 10424 10424
DOMAIN 10425 10425
DOMAIN 10426 10426
DOMAIN 10427 10427
DOMAIN 10428 10428
DOMAIN 10429 10429
DOMAIN 10430 10430
DOMAIN 10431 10431
DOMAIN 10432 10432
DOMAIN 10433 10433
DOMAIN 10434 10434
DOMAIN 10435 10435
DOMAIN 10436 10436
DOMAIN 10437 10437
DOMAIN 10438 10438
DOMAIN 10439 10439
DOMAIN 10440 10440
DOMAIN 10441 10441
DOMAIN 10442 10442
DOMAIN 10443 10443
DOMAIN 10444 10444
DOMAIN 10445 10445
DOMAIN 10446 10446
DOMAIN 10447 10447
DOMAIN 10448 10448
DOMAIN 10449 10449
DOMAIN 10450 10450
DOMAIN 10451 10451
DOMAIN 10452 10452
DOMAIN 10453 10453
DOMAIN 10454 10454
DOMAIN 10455 10455
DOMAIN 10456 10456
DOMAIN 10457 10457
DOMAIN 10458 10458
DOMAIN 10459 10459
DOMAIN 10460 10460
DOMAIN 10461 10461
DOMAIN 10462 10462
DOMAIN 10463 10463
DOMAIN 10464 10464
DOMAIN 10465 10465
DOMAIN 10466 10466
DOMAIN 10467 10467
DOMAIN 10468 10468
DOMAIN 10469 10469
DOMAIN 10470 10470
DOMAIN 10471 10471
DOMAIN 10472 10472
DOMAIN 10473 10473
DOMAIN 10474 10474
DOMAIN 10475 10475
DOMAIN 10476 10476
DOMAIN 10477 10477
DOMAIN 10478 10478
DOMAIN 10479 10479
DOMAIN 10480 10480
DOMAIN 10481 10481
DOMAIN 10482 10482
DOMAIN 10483 10483
DOMAIN 10484 10484
DOMAIN 10485 10485
DOMAIN 10486 10486
DOMAIN 10487 10487
DOMAIN 10488 10488
DOMAIN 10489 10489
DOMAIN 10490 10490
DOMAIN 10491 10491
DOMAIN 10492 10492
DOMAIN 10493 10493
DOMAIN 10494 10494
DOMAIN 10495 10495
DOMAIN 10496 10496
DOMAIN 10497 10497
DOMAIN 10498 10498
DOMAIN 10499 10499
DOMAIN 10500 10500
DOMAIN 10501 10501
DOMAIN 10502 10502
DOMAIN 10503 10503
DOMAIN 10504 10504
DOMAIN 10505 10505
DOMAIN 10506 10506
DOMAIN 10507 10507
DOMAIN 10508 10508
DOMAIN 10509 10509
DOMAIN 10510 10510
DOMAIN 10511 10511
DOMAIN 10512 10512
DOMAIN 10513 10513
DOMAIN 10514 10514
DOMAIN 10515 10515
DOMAIN 10516 10516
DOMAIN 10517 10517
DOMAIN 10518 10518
DOMAIN 10519 10519
DOMAIN 10520 10520
DOMAIN 10521 10521
DOMAIN 10522 10522
DOMAIN 10523 10523
DOMAIN 10524 10524
DOMAIN 10525 10525
DOMAIN 10526 10526
DOMAIN 10527 10527
DOMAIN 10528 10528
DOMAIN 10529 10529
DOMAIN 10530 10530
DOMAIN 10531 10531
DOMAIN 10532 10532
DOMAIN 10533 10533
DOMAIN 10534 10534
DOMAIN 10535 10535
DOMAIN 10536 10536
DOMAIN 10537 10537
DOMAIN 10538 10538
DOMAIN 10539 10539
DOMAIN 10540 10540
DOMAIN 10541 10541
DOMAIN 10542 10542
DOMAIN 10543 10543
DOMAIN 10544 10544
DOMAIN 10545 10545
DOMAIN 10546 10546
DOMAIN 10547 10547
DOMAIN 10548 10548
DOMAIN 10549 10549
DOMAIN 10550 10550
DOMAIN 10551 10551
DOMAIN 10552 10552
DOMAIN 10553 10553
DOMAIN 10554 10554
DOMAIN 10555 10555
DOMAIN 10556 10556
DOMAIN 10557 10557
DOMAIN 10558 10558
DOMAIN 10559 10559
DOMAIN 10560 10560
DOMAIN 10561 10561
DOMAIN 10562 10562
DOMAIN 10563 10563
DOMAIN 10564 10564
DOMAIN 10565 10565
DOMAIN 10566 10566
DOMAIN 10567 10567
DOMAIN 10568 10568
DOMAIN 10569 10569
DOMAIN 10570 10570
DOMAIN 10571 10571
DOMAIN 10572 10572
DOMAIN 10573 10573
DOMAIN 10574 10574
DOMAIN 10575 10575
DOMAIN 10576 10576
DOMAIN 10577 10577
DOMAIN 10578 10578
DOMAIN 105

| Query Match | Best Local Similarity | Matches 123; | Conservative | Score 357; | DB 1; | Length 1221; |
|---|-----------------------|--------------|-------------------------|------------|--------|--------------------|
| 154 EEOGCEVKAD-GFLICEFHFPATC-----RPLAV-----EPGAAAAVS-----IT 192 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| 521 EGOGCEBNPLGLGPCN-HYMLSCGEGEPLIYEVRRRPEEARPRVRSBEMASREALS 579 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| 193 YGT----PFAAGADFO---ALP-----VGSSAAVAPLGLOLM-----CTAPPG 229 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| 580 LGTAELPNSLPGDDDDDECLMLPGELCOHICINTVGSRCACFPGFELQDGRGCRPDRG 639 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| 230 AVQGHMARE-APAPMCCSV-----ENGGCEHMAONALPGAPRPOCGAGAA 272 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| APQJLDTRESAPRSESNQVSPNTIPLRPQNTCKDNBPCQKQVCRRVGGDTAMSCSPRGTA 699 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| 273 LQADGRSC-----TASATQSCNDLCEHFCFPPNDPQGSYSC-----MCEGYRIADQHR 322 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| 700 IMADGVSCEDQDDCLMGTHDCS--WKQFCV---NLTGSGFVUNHTVLCAAGYILMA-HRK 753 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| 323 CEDVDCLILEPSSC--PQRCVNTQGGFECH---CYPTIYLDVDECEVPYDPGFRA--NC 374 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| 754 CVDINCEVTLMLCTRAEHCVTNTPGSGFOCYKALTCPEGYVLTDEGCH-VDECVTGTNHC 812 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| 375 E--VOCOPLMONTYILV---CAEGRAPITPH-----EPRRCMFCNOTACPA 415 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| 813 QAGSFCON-TKGSFYCOAQRRCMDGFLDQPBENCNCVDINECTSLLEPCRSQSFSCINTVGSY 871 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |
| 416 DCDPNTQASCECPBGV-IIDDGFTCTDDCENGGFCSG---VGHNLPGFECTICGP 468 | 1221 AA; | 131818 MW; | 87DB2A10A8FDC45F CRC64; | 12.28; | 29.58; | Pred. No. 2.3e-16; |

| | |
|------|--|
| DR | PROSITE: PS50088; ANK_REPEAT: 5. |
| DR | PROSITE: PS50297; ANK_REPEAT_REGION: 1. |
| DR | PROSITE: PS00010; ASX_HYDROXYL: 11. |
| DR | PROSITE: PS00022; EGF_1; 28. |
| DR | PROSITE: PS01186; EGF_2; 21. |
| DR | PROSITE: PS01187; EGF_CA: 9. |
| KW | differentiation; Neurogenesis; Repeat; EGF-like domain; Transmembrane; Glycoprotein; Proto-oncogene; ANK repeat; Signal. |
| FT | SIGNAL 1 20 |
| FT | CHAIN 1 1964 |
| FT | DOMAIN 21 1443 |
| FT | TRANSMEM 1444 1464 |
| FT | DOMAIN 1465 1964 |
| FT | DOMAIN 21 60 |
| FT | DOMAIN 61 112 |
| FT | DOMAIN 115 152 |
| FT | DOMAIN 153 189 |
| FT | DOMAIN 191 229 |
| FT | DOMAIN 231 271 |
| FT | DOMAIN 273 309 |
| FT | DOMAIN 311 350 |
| FT | DOMAIN 352 388 |
| FT | DOMAIN 389 427 |
| FT | DOMAIN 429 470 |
| FT | DOMAIN 472 508 |
| FT | DOMAIN 510 546 |
| FT | DOMAIN 548 584 |
| FT | DOMAIN 586 622 |
| FT | DOMAIN 622 656 |
| FT | DOMAIN 658 686 |
| FT | DOMAIN 688 724 |
| FT | DOMAIN 726 762 |
| FT | DOMAIN 764 800 |
| FT | DOMAIN 803 839 |
| FT | DOMAIN 841 877 |
| FT | DOMAIN 878 924 |
| FT | DOMAIN 926 962 |
| FT | DOMAIN 964 1000 |
| FT | DOMAIN 1002 1040 |
| FT | DOMAIN 1042 1081 |
| FT | DOMAIN 1083 1122 |
| FT | DOMAIN 1126 1167 |
| FT | DOMAIN 1168 1208 |
| FT | REPEAT 1209 1242 |
| FT | REPEAT 1243 1282 |
| FT | REPEAT 1283 1657 |
| FT | REPEAT 1658 1691 |
| FT | REPEAT 1691 1724 |
| FT | REPEAT 1724 1757 |
| FT | REPEAT 1757 1790 |
| FT | DISULEFD 25 38 |
| FT | DISULEFD 32 48 |
| FT | DISULEFD 50 59 |
| FT | DISULEFD 65 77 |
| FT | DISULEFD 71 100 |
| FT | DISULEFD 102 111 |
| FT | DISULEFD 119 130 |
| FT | DISULEFD 124 140 |
| FT | DISULEFD 142 151 |
| FT | DISULEFD 157 168 |
| FT | DISULEFD 162 177 |
| FT | DISULEFD 179 188 |
| FT | DISULEFD 195 208 |
| FT | DISULEFD 202 217 |
| FT | DISULEFD 219 228 |
| FT | DISULEFD 235 246 |
| FT | DISULEFD 240 259 |
| FT | DISULEFD 261 270 |
| FT | DISULEFD 277 288 |
| FT | DISULEFD 282 297 |
| FT | DISULEFD 299 308 |
| FT | DISULEFD 315 329 |
| FT | DISULEFD 323 338 |
| FT | DISULEFD 338 349 |
| FT | DISULEFD 349 360 |
| FT | DISULEFD 360 371 |
| FT | DISULEFD 371 382 |
| FT | DISULEFD 382 393 |
| FT | DISULEFD 393 404 |
| FT | DISULEFD 404 415 |
| FT | DISULEFD 415 426 |
| FT | DISULEFD 426 437 |
| FT | DISULEFD 437 448 |
| FT | DISULEFD 448 459 |
| FT | DISULEFD 459 470 |
| FT | DISULEFD 470 481 |
| FT | DISULEFD 481 492 |
| FT | DISULEFD 492 503 |
| FT | DISULEFD 503 514 |
| FT | DISULEFD 514 525 |
| FT | DISULEFD 525 536 |
| FT | DISULEFD 536 547 |
| FT | DISULEFD 547 558 |
| FT | DISULEFD 558 569 |
| FT | DISULEFD 569 580 |
| FT | DISULEFD 580 591 |
| FT | DISULEFD 591 602 |
| FT | DISULEFD 602 613 |
| FT | DISULEFD 613 624 |
| FT | DISULEFD 624 635 |
| FT | DISULEFD 635 646 |
| FT | DISULEFD 646 657 |
| FT | DISULEFD 657 668 |
| FT | DISULEFD 668 679 |
| FT | DISULEFD 679 690 |
| FT | DISULEFD 690 701 |
| FT | DISULEFD 701 712 |
| FT | DISULEFD 712 723 |
| FT | DISULEFD 723 734 |
| FT | DISULEFD 734 745 |
| FT | DISULEFD 745 756 |
| FT | DISULEFD 756 767 |
| FT | DISULEFD 767 778 |
| FT | DISULEFD 778 789 |
| FT | DISULEFD 789 800 |
| FT | DISULEFD 800 811 |
| FT | DISULEFD 811 822 |
| FT | DISULEFD 822 833 |
| FT | DISULEFD 833 844 |
| FT | DISULEFD 844 855 |
| FT | DISULEFD 855 866 |
| FT | DISULEFD 866 877 |
| FT | DISULEFD 877 888 |
| FT | DISULEFD 888 899 |
| FT | DISULEFD 899 910 |
| FT | DISULEFD 910 921 |
| FT | DISULEFD 921 932 |
| FT | DISULEFD 932 943 |
| FT | DISULEFD 943 954 |
| FT | DISULEFD 954 965 |
| FT | DISULEFD 965 976 |
| FT | DISULEFD 976 987 |
| FT | DISULEFD 987 998 |
| FT | DISULEFD 998 1009 |
| FT | DISULEFD 1009 1020 |
| FT | DISULEFD 1020 1031 |
| FT | DISULEFD 1031 1042 |
| FT | DISULEFD 1042 1053 |
| FT | DISULEFD 1053 1064 |
| FT | DISULEFD 1064 1075 |
| FT | DISULEFD 1075 1086 |
| FT | DISULEFD 1086 1097 |
| FT | DISULEFD 1097 1108 |
| FT | DISULEFD 1108 1119 |
| FT | DISULEFD 1119 1130 |
| FT | DISULEFD 1130 1141 |
| FT | DISULEFD 1141 1152 |
| FT | DISULEFD 1152 1163 |
| FT | DISULEFD 1163 1174 |
| FT | DISULEFD 1174 1185 |
| FT | DISULEFD 1185 1196 |
| FT</ | |

```

FT DISULFID 340 349 BY SIMILARITY.
FT DISULFID 356 376 BY SIMILARITY.
FT DISULFID 361 376 BY SIMILARITY.
FT DISULFID 378 387 BY SIMILARITY.
FT DISULFID 393 404 BY SIMILARITY.
FT DISULFID 398 415 BY SIMILARITY.
FT DISULFID 417 426 BY SIMILARITY.
FT DISULFID 433 449 BY SIMILARITY.
FT DISULFID 443 458 BY SIMILARITY.
FT DISULFID 460 469 BY SIMILARITY.
FT DISULFID 476 487 BY SIMILARITY.
FT DISULFID 481 496 BY SIMILARITY.
FT DISULFID 498 507 BY SIMILARITY.
FT DISULFID 514 525 BY SIMILARITY.
FT DISULFID 519 534 BY SIMILARITY.
FT DISULFID 536 545 BY SIMILARITY.
FT DISULFID 552 563 BY SIMILARITY.
FT DISULFID 557 572 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
FT DISULFID 590 601 BY SIMILARITY.
FT DISULFID 595 610 BY SIMILARITY.
FT DISULFID 612 621 BY SIMILARITY.
FT DISULFID 626 637 BY SIMILARITY.
FT DISULFID 631 646 BY SIMILARITY.
FT DISULFID 648 655 BY SIMILARITY.
FT DISULFID 662 669 BY SIMILARITY.
FT DISULFID 664 674 BY SIMILARITY.
FT DISULFID 676 685 BY SIMILARITY.
FT DISULFID 692 703 BY SIMILARITY.
FT DISULFID 712 723 BY SIMILARITY.
FT DISULFID 714 723 BY SIMILARITY.
FT DISULFID 730 741 BY SIMILARITY.
FT DISULFID 735 750 BY SIMILARITY.
FT DISULFID 752 761 BY SIMILARITY.
FT DISULFID 768 779 BY SIMILARITY.
FT DISULFID 773 788 BY SIMILARITY.
FT DISULFID 790 799 BY SIMILARITY.
FT DISULFID 807 818 BY SIMILARITY.
FT DISULFID 812 827 BY SIMILARITY.
FT DISULFID 829 838 BY SIMILARITY.
FT DISULFID 845 856 BY SIMILARITY.
FT DISULFID 850 865 BY SIMILARITY.
FT DISULFID 867 876 BY SIMILARITY.
FT DISULFID 882 903 BY SIMILARITY.
FT DISULFID 897 912 BY SIMILARITY.
FT DISULFID 914 923 BY SIMILARITY.
FT DISULFID 930 941 BY SIMILARITY.
FT DISULFID 935 950 BY SIMILARITY.
FT DISULFID 952 961 BY SIMILARITY.
FT DISULFID 968 979 BY SIMILARITY.
FT DISULFID 973 988 BY SIMILARITY.
FT DISULFID 990 999 BY SIMILARITY.
FT DISULFID 1006 1019 BY SIMILARITY.
FT DISULFID 1011 1028 BY SIMILARITY.
FT DISULFID 1030 1039 BY SIMILARITY.
FT DISULFID 1046 1057 BY SIMILARITY.
FT DISULFID 1051 1069 BY SIMILARITY.
FT DISULFID 1071 1080 BY SIMILARITY.
FT DISULFID 1087 1098 BY SIMILARITY.

```

Query Match 11.3%; Score 329.5; DB 1; Length 1964;
 Best Local Similarity 26.3%; Pred. No. 2.2e-14;
 Matches 128; Conservative 33; Mismatches 143; Indels 183; Gaps 31;

```

DB 134 GFLVAVSAEATVPSSEPTMEQOCEVADFLCE-FHPATCRPLAVPGAAAAVSIT 192
DB 35 GGTCLMLSGGIC-----GC-----APGLGTGCPDPDC----- 66
OY 193 YCTPFAARGADFOAL---FVGSSAAVAPLGLMCTAPPG---AVQGHMAREAPGAMDC 245
DB 67 -DTOLCKNGSCQALLPFPSSRSPPLTPHFSCCTGSGFTGDCQTHLELCPSPF-C 124
OY 246 SVENGCERHACNAIRPGAPRCQCPAGALQADGRSCTASATOSCNLDCEHFCVNPDPG- 304

```

```

DB 125 S---NGG--HCYVQASGRPOCCCEPGWT-----GEQO-----LRDFCSANCANMG 166
OY 305 -----SYSCMETGYRLAADQHRCE-DVDDCIIEPSPQ--RCVNTGGFECCHC--- 351
DB 167 VCLATVPQICRCRPPGF-----EGHTCERDINECELEPGPCPGOTSGHNTLSGYQCLCPVG 222
OY 352 --TPNDIVDGEVEVDPFRANCEYOCQPL---NOTSYLCVAGAPRIPIHE----- 400
DB 223 QEGPOCKLRGAC--PGSCLNG--TCOLYBEGSHFHLCLCPGFGGLCEMNPDC 277
OY 401 -PHRCMFQNTACPADCPNTOASCEPEGYLDGFCIT--DIDCE-----NGGF 450
DB 278 VRHQCO---NGATCLDGLDIT--CLCRXTW---KGMDCSEDDIDCEARGPFRNGST 328
OY 451 C-----SGVCHNLPGEFCICGP----- 468
DB 329 COMTAGSFHCVCVSGMGAGCEENLDCAATCAPGSTCIDRVGSFSCLPGRGTLCH 388
OY 469 --DSALVR--HTTDDSGKVDG-----GDSGS-----GEPSP----- 499
DB 389 LEDMCLSQPCRVNAQSTFPLTGSTYLICQPGYSGSTCHQDDECOMAQGSPCEHGS 448
OY 500 ---TPGS 503
DB 449 CINTPGS 455

```

RESULT 14
 MTN2_MOUSE STANDARD; PRT; 956 AA.
 AC 008746;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE MATRILIN-2 PRECURSOR.
 GN MATN2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-Limb;
 RX MEDLINE=97238663; PubMed=9083061;
 RA Deak F., Plecha D., Bachnati C., Paulsson M., Kiss I.;
 RT "Primary structure and expression of matrilin-2, the closest relative
 of cartilage matrix protein within the von Willebrand factor type A-
 RT like module superfamily";
 RL J. Biol. Chem. 272:9268-9274(1997).
 CC -I- FUNCTION: INVOLVED IN MATRIX ASSEMBLY (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: SECRETED.
 CC -I- TISSUE SPECIFICITY: DETECTED IN A VARIETY OF ORGANS, INCLUDING
 CC CALVARIA, UTERUS, HEART, AND BRAIN, AS WELL AS FIBROBLAST AND
 CC OSTEOBLAST CELL LINES.
 CC -I- SIMILARITY: CONTAINS 10 EGF-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 VMFA DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 CC EMBL; 069262; AAC53163.1; -
 CC HSSP; P07204; 1ZAQ.
 CC MGD; MGI:109613; Matn2.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR002035; VMFA.

DR Pfam: PF00008; EGF: 10.
 DR Pfam: PF00092; vwa: 2.
 DR PRINTS: PR00453; VWFADOMAIN.
 DR SMART: SM00181; EGF: 10.
 DR SMART: SM00327; vwa: 2.
 DR PROSITE: PS00010; ASX_HYDROXYL: 7.
 DR PROSITE: PS01186; EGF_2: 9.
 DR PROSITE: PS0234; vwa: 2.
 DR EGF-like domain; signal; Glycoprotein; Repeat; Coiled coil.
 KW EGF-like domain; signal; Glycoprotein; Repeat; Coiled coil.
 FT CHAIN 1 23
 FT CHAIN 24 956
 FT CHAIN 57 232
 FT CHAIN 238 278
 FT CHAIN 279 319
 FT CHAIN 320 360
 FT CHAIN 361 401
 FT CHAIN 402 442
 FT CHAIN 443 483
 FT CHAIN 484 524
 FT CHAIN 525 565
 FT CHAIN 566 606
 FT CHAIN 607 647
 FT CHAIN 647 687
 FT CHAIN 687 727
 FT CHAIN 727 767
 FT CHAIN 767 807
 FT CHAIN 807 847
 FT CHAIN 847 887
 FT CHAIN 887 927
 FT CHAIN 927 967
 FT CHAIN 967 1007
 FT CHAIN 1007 1047
 FT CHAIN 1047 1087
 FT CHAIN 1087 1127
 FT CHAIN 1127 1167
 FT CHAIN 1167 1207
 FT CHAIN 1207 1247
 FT CHAIN 1247 1287
 FT CHAIN 1287 1327
 FT CHAIN 1327 1367
 FT CHAIN 1367 1407
 FT CHAIN 1407 1447
 FT CHAIN 1447 1487
 FT CHAIN 1487 1527
 FT CHAIN 1527 1567
 FT CHAIN 1567 1607
 FT CHAIN 1607 1647
 FT CHAIN 1647 1687
 FT CHAIN 1687 1727
 FT CHAIN 1727 1767
 FT CHAIN 1767 1807
 FT CHAIN 1807 1847
 FT CHAIN 1847 1887
 FT CHAIN 1887 1927
 FT CHAIN 1927 1967
 FT CHAIN 1967 2007
 FT CHAIN 2007 2047
 FT CHAIN 2047 2087
 FT CHAIN 2087 2127
 FT CHAIN 2127 2167
 FT CHAIN 2167 2207
 FT CHAIN 2207 2247
 FT CHAIN 2247 2287
 FT CHAIN 2287 2327
 FT CHAIN 2327 2367
 FT CHAIN 2367 2407
 FT CHAIN 2407 2447
 FT CHAIN 2447 2487
 FT CHAIN 2487 2527
 FT CHAIN 2527 2567
 FT CHAIN 2567 2607
 FT CHAIN 2607 2647
 FT CHAIN 2647 2687
 FT CHAIN 2687 2727
 FT CHAIN 2727 2767
 FT CHAIN 2767 2807
 FT CHAIN 2807 2847
 FT CHAIN 2847 2887
 FT CHAIN 2887 2927
 FT CHAIN 2927 2967
 FT CHAIN 2967 3007
 FT CHAIN 3007 3047
 FT CHAIN 3047 3087
 FT CHAIN 3087 3127
 FT CHAIN 3127 3167
 FT CHAIN 3167 3207
 FT CHAIN 3207 3247
 FT CHAIN 3247 3287
 FT CHAIN 3287 3327
 FT CHAIN 3327 3367
 FT CHAIN 3367 3407
 FT CHAIN 3407 3447
 FT CHAIN 3447 3487
 FT CHAIN 3487 3527
 FT CHAIN 3527 3567
 FT CHAIN 3567 3607
 FT CHAIN 3607 3647
 FT CHAIN 3647 3687
 FT CHAIN 3687 3727
 FT CHAIN 3727 3767
 FT CHAIN 3767 3807
 FT CHAIN 3807 3847
 FT CHAIN 3847 3887
 FT CHAIN 3887 3927
 FT CHAIN 3927 3967
 FT CHAIN 3967 4007
 FT CHAIN 4007 4047
 FT CHAIN 4047 4087
 FT CHAIN 4087 4127
 FT CHAIN 4127 4167
 FT CHAIN 4167 4207
 FT CHAIN 4207 4247
 FT CHAIN 4247 4287
 FT CHAIN 4287 4327
 FT CHAIN 4327 4367
 FT CHAIN 4367 4407
 FT CHAIN 4407 4447
 FT CHAIN 4447 4487
 FT CHAIN 4487 4527
 FT CHAIN 4527 4567
 FT CHAIN 4567 4607
 FT CHAIN 4607 4647
 FT CHAIN 4647 4687
 FT CHAIN 4687 4727
 FT CHAIN 4727 4767
 FT CHAIN 4767 4807
 FT CHAIN 4807 4847
 FT CHAIN 4847 4887
 FT CHAIN 4887 4927
 FT CHAIN 4927 4967
 FT CHAIN 4967 5007
 FT CHAIN 5007 5047
 FT CHAIN 5047 5087
 FT CHAIN 5087 5127
 FT CHAIN 5127 5167
 FT CHAIN 5167 5207
 FT CHAIN 5207 5247
 FT CHAIN 5247 5287
 FT CHAIN 5287 5327
 FT CHAIN 5327 5367
 FT CHAIN 5367 5407
 FT CHAIN 5407 5447
 FT CHAIN 5447 5487
 FT CHAIN 5487 5527
 FT CHAIN 5527 5567
 FT CHAIN 5567 5607
 FT CHAIN 5607 5647
 FT CHAIN 5647 5687
 FT CHAIN 5687 5727
 FT CHAIN 5727 5767
 FT CHAIN 5767 5807
 FT CHAIN 5807 5847
 FT CHAIN 5847 5887
 FT CHAIN 5887 5927
 FT CHAIN 5927 5967
 FT CHAIN 5967 6007
 FT CHAIN 6007 6047
 FT CHAIN 6047 6087
 FT CHAIN 6087 6127
 FT CHAIN 6127 6167
 FT CHAIN 6167 6207
 FT CHAIN 6207 6247
 FT CHAIN 6247 6287
 FT CHAIN 6287 6327
 FT CHAIN 6327 6367
 FT CHAIN 6367 6407
 FT CHAIN 6407 6447
 FT CHAIN 6447 6487
 FT CHAIN 6487 6527
 FT CHAIN 6527 6567
 FT CHAIN 6567 6607
 FT CHAIN 6607 6647
 FT CHAIN 6647 6687
 FT CHAIN 6687 6727
 FT CHAIN 6727 6767
 FT CHAIN 6767 6807
 FT CHAIN 6807 6847
 FT CHAIN 6847 6887
 FT CHAIN 6887 6927
 FT CHAIN 6927 6967
 FT CHAIN 6967 7007
 FT CHAIN 7007 7047
 FT CHAIN 7047 7087
 FT CHAIN 7087 7127
 FT CHAIN 7127 7167
 FT CHAIN 7167 7207
 FT CHAIN 7207 7247
 FT CHAIN 7247 7287
 FT CHAIN 7287 7327
 FT CHAIN 7327 7367
 FT CHAIN 7367 7407
 FT CHAIN 7407 7447
 FT CHAIN 7447 7487
 FT CHAIN 7487 7527
 FT CHAIN 7527 7567
 FT CHAIN 7567 7607
 FT CHAIN 7607 7647
 FT CHAIN 7647 7687
 FT CHAIN 7687 7727
 FT CHAIN 7727 7767
 FT CHAIN 7767 7807
 FT CHAIN 7807 7847
 FT CHAIN 7847 7887
 FT CHAIN 7887 7927
 FT CHAIN 7927 7967
 FT CHAIN 7967 8007
 FT CHAIN 8007 8047
 FT CHAIN 8047 8087
 FT CHAIN 8087 8127
 FT CHAIN 8127 8167
 FT CHAIN 8167 8207
 FT CHAIN 8207 8247
 FT CHAIN 8247 8287
 FT CHAIN 8287 8327
 FT CHAIN 8327 8367
 FT CHAIN 8367 8407
 FT CHAIN 8407 8447
 FT CHAIN 8447 8487
 FT CHAIN 8487 8527
 FT CHAIN 8527 8567
 FT CHAIN 8567 8607
 FT CHAIN 8607 8647
 FT CHAIN 8647 8687
 FT CHAIN 8687 8727
 FT CHAIN 8727 8767
 FT CHAIN 8767 8807
 FT CHAIN 8807 8847
 FT CHAIN 8847 8887
 FT CHAIN 8887 8927
 FT CHAIN 8927 8967
 FT CHAIN 8967 9007
 FT CHAIN 9007 9047
 FT CHAIN 9047 9087
 FT CHAIN 9087 9127
 FT CHAIN 9127 9167
 FT CHAIN 9167 9207
 FT CHAIN 9207 9247
 FT CHAIN 9247 9287
 FT CHAIN 9287 9327
 FT CHAIN 9327 9367
 FT CHAIN 9367 9407
 FT CHAIN 9407 9447
 FT CHAIN 9447 9487
 FT CHAIN 9487 9527
 FT CHAIN 9527 9567
 FT CHAIN 9567 9607
 FT CHAIN 9607 9647
 FT CHAIN 9647 9687
 FT CHAIN 9687 9727
 FT CHAIN 9727 9767
 FT CHAIN 9767 9807
 FT CHAIN 9807 9847
 FT CHAIN 9847 9887
 FT CHAIN 9887 9927
 FT CHAIN 9927 9967
 FT CHAIN 9967 10007
 FT CHAIN 10007 10047
 FT CHAIN 10047 10087
 FT CHAIN 10087 10127
 FT CHAIN 10127 10167
 FT CHAIN 10167 10207
 FT CHAIN 10207 10247
 FT CHAIN 10247 10287
 FT CHAIN 10287 10327
 FT CHAIN 10327 10367
 FT CHAIN 10367 10407
 FT CHAIN 10407 10447
 FT CHAIN 10447 10487
 FT CHAIN 10487 10527
 FT CHAIN 10527 10567
 FT CHAIN 10567 10607
 FT CHAIN 10607 10647
 FT CHAIN 10647 10687
 FT CHAIN 10687 10727
 FT CHAIN 10727 10767
 FT CHAIN 10767 10807
 FT CHAIN 10807 10847
 FT CHAIN 10847 10887
 FT CHAIN 10887 10927
 FT CHAIN 10927 10967
 FT CHAIN 10967 11007
 FT CHAIN 11007 11047
 FT CHAIN 11047 11087
 FT CHAIN 11087 11127
 FT CHAIN 11127 11167
 FT CHAIN 11167 11207
 FT CHAIN 11207 11247
 FT CHAIN 11247 11287
 FT CHAIN 11287 11327
 FT CHAIN 11327 11367
 FT CHAIN 11367 11407
 FT CHAIN 11407 11447
 FT CHAIN 11447 11487
 FT CHAIN 11487 11527
 FT CHAIN 11527 11567
 FT CHAIN 11567 11607
 FT CHAIN 11607 11647
 FT CHAIN 11647 11687
 FT CHAIN 11687 11727
 FT CHAIN 11727 11767
 FT CHAIN 11767 11807
 FT CHAIN 11807 11847
 FT CHAIN 11847 11887
 FT CHAIN 11887 11927
 FT CHAIN 11927 11967
 FT CHAIN 11967 12007
 FT CHAIN 12007 12047
 FT CHAIN 12047 12087
 FT CHAIN 12087 12127
 FT CHAIN 12127 12167
 FT CHAIN 12167 12207
 FT CHAIN 12207 12247
 FT CHAIN 12247 12287
 FT CHAIN 12287 12327
 FT CHAIN 12327 12367
 FT CHAIN 12367 12407
 FT CHAIN 12407 12447
 FT CHAIN 12447 12487
 FT CHAIN 12487 12527
 FT CHAIN 12527 12567
 FT CHAIN 12567 12607
 FT CHAIN 12607 12647
 FT CHAIN 12647 12687
 FT CHAIN 12687 12727
 FT CHAIN 12727 12767
 FT CHAIN 12767 12807
 FT CHAIN 12807 12847
 FT CHAIN 12847 12887
 FT CHAIN 12887 12927
 FT CHAIN 12927 12967
 FT CHAIN 12967 13007
 FT CHAIN 13007 13047
 FT CHAIN 13047 13087
 FT CHAIN 13087 13127
 FT CHAIN 13127 13167
 FT CHAIN 13167 13207
 FT CHAIN 13207 13247
 FT CHAIN 13247 13287
 FT CHAIN 13287 13327
 FT CHAIN 13327 13367
 FT CHAIN 13367 13407
 FT CHAIN 13407 13447
 FT CHAIN 13447 13487
 FT CHAIN 13487 13527
 FT CHAIN 13527 13567
 FT CHAIN 13567 13607
 FT CHAIN 13607 13647
 FT CHAIN 13647 13687
 FT CHAIN 13687 13727
 FT CHAIN 13727 13767
 FT CHAIN 13767 13807
 FT CHAIN 13807 13847
 FT CHAIN 13847 13887
 FT CHAIN 13887 13927
 FT CHAIN 13927 13967
 FT CHAIN 13967 14007
 FT CHAIN 14007 14047
 FT CHAIN 14047 14087
 FT CHAIN 14087 14127
 FT CHAIN 14127 14167
 FT CHAIN 14167 14207
 FT CHAIN 14207 14247
 FT CHAIN 14247 14287
 FT CHAIN 14287 14327
 FT CHAIN 14327 14367
 FT CHAIN 14367 14407
 FT CHAIN 14407 14447
 FT CHAIN 14447 14487
 FT CHAIN 14487 14527
 FT CHAIN 14527 14567
 FT CHAIN 14567 14607
 FT CHAIN 14607 14647
 FT CHAIN 14647 14687
 FT CHAIN 14687 14727
 FT CHAIN 14727 14767
 FT CHAIN 14767 14807
 FT CHAIN 14807 14847
 FT CHAIN 14847 14887
 FT CHAIN 14887 14927
 FT CHAIN 14927 14967
 FT CHAIN 14967 15007
 FT CHAIN 15007 15047
 FT CHAIN 15047 15087
 FT CHAIN 15087 15127
 FT CHAIN 15127 15167
 FT CHAIN 15167 15207
 FT CHAIN 15207 15247
 FT CHAIN 15247 15287
 FT CHAIN 15287 15327
 FT CHAIN 15327 15367
 FT CHAIN 15367 15407
 FT CHAIN 15407 15447
 FT CHAIN 15447 15487
 FT CHAIN 15487 15527
 FT CHAIN 15527 15567
 FT CHAIN 15567 15607
 FT CHAIN 15607 15647
 FT CHAIN 15647 15687
 FT CHAIN 15687 15727
 FT CHAIN 15727 15767
 FT CHAIN 15767 15807
 FT CHAIN 15807 15847
 FT CHAIN 15847 15887
 FT CHAIN 15887 15927
 FT CHAIN 15927 15967
 FT CHAIN 15967 16007
 FT CHAIN 16007 16047
 FT CHAIN 16047 16087
 FT CHAIN 16087 16127
 FT CHAIN 16127 16167
 FT CHAIN 16167 16207
 FT CHAIN 16207 16247
 FT CHAIN 16247 16287
 FT CHAIN 16287 16327
 FT CHAIN 16327 16367
 FT CHAIN 16367 16407
 FT CHAIN 16407 16447
 FT CHAIN 16447 16487
 FT CHAIN 16487 16527
 FT CHAIN 16527 16567
 FT CHAIN 16567 16607
 FT CHAIN 16607 16647
 FT CHAIN 16647 16687
 FT CHAIN 16687 16727
 FT CHAIN 16727 16767
 FT CHAIN 16767 16807
 FT CHAIN 16807 16847
 FT CHAIN 16847 16887
 FT CHAIN 16887 16927
 FT CHAIN 16927 16967
 FT CHAIN 16967 17007
 FT CHAIN 17007 17047
 FT CHAIN 17047 17087
 FT CHAIN 17087 17127
 FT CHAIN 17127 17167
 FT CHAIN 17167 17207
 FT CHAIN 17207 17247
 FT CHAIN 17247 17287
 FT CHAIN 17287 17327
 FT CHAIN 17327 17367
 FT CHAIN 17367 17407
 FT CHAIN 17407 17447
 FT CHAIN 17447 17487
 FT CHAIN 17487 17527
 FT CHAIN 17527 17567
 FT CHAIN 17567 17607
 FT CHAIN 17607 17647
 FT CHAIN 17647 17687
 FT CHAIN 17687 17727
 FT CHAIN 17727 17767
 FT CHAIN 17767 17807
 FT CHAIN 17807 17847
 FT CHAIN 17847 17887
 FT CHAIN 17887 17927
 FT CHAIN 17927 17967
 FT CHAIN 17967 18007
 FT CHAIN 18007 18047
 FT CHAIN 18047 18087
 FT CHAIN 18087 18127
 FT CHAIN 18127 18167
 FT CHAIN 18167 18207
 FT CHAIN 18207 18247
 FT CHAIN 18247 18287
 FT CHAIN 18287 18327
 FT CHAIN 18327 18367
 FT CHAIN 18367 18407
 FT CHAIN 18407 18447
 FT CHAIN 18447 18487
 FT CHAIN 18487 18527
 FT CHAIN 18527 18567
 FT CHAIN 18567 18607
 FT CHAIN 18607 18647
 FT CHAIN 18647 18687
 FT CHAIN 18687 18727
 FT CHAIN 18727 18767
 FT CHAIN 18767 18807
 FT CHAIN 18807 18847
 FT CHAIN 18847 18887
 FT CHAIN 18887 18927
 FT CHAIN 18927 18967
 FT CHAIN 18967 19007
 FT CHAIN 19007 19047
 FT CHAIN 19047 19087
 FT CHAIN 19087 19127
 FT CHAIN 19127 19167
 FT CHAIN 19167 19207
 FT CHAIN 19207 19247
 FT CHAIN 19247 19287
 FT CHAIN 19287 19327
 FT CHAIN 19327 19367
 FT CHAIN 19367 19407
 FT CHAIN 19407 19447
 FT CHAIN 19447 19487
 FT CHAIN 19487 19527
 FT CHAIN 19527 19567
 FT CHAIN 19567 19607
 FT CHAIN 19607 19647
 FT CHAIN 19647 19687
 FT CHAIN 19687 19727
 FT CHAIN 19727 19767
 FT CHAIN 19767 19807
 FT CHAIN 19807 19847
 FT CHAIN 19847 19887
 FT CHAIN 19887 19927
 FT CHAIN 19927 19967
 FT CHAIN 19967 20007
 FT CHAIN 20007 20047
 FT CHAIN 20047 20087
 FT CHAIN 20087 20127
 FT CHAIN 20127 20167
 FT CHAIN 20167 20207
 FT CHAIN 20207 20247
 FT CHAIN 20247 20287
 FT CHAIN 20287 20327
 FT CHAIN 20327 20367
 FT CHAIN 20367 20407
 FT CHAIN 20407 20447
 FT CHAIN 20447 20487
 FT CHAIN 20487 20527
 FT CHAIN 20527 20567
 FT CHAIN 20567 20607
 FT CHAIN 20607 20647
 FT CHAIN 20647 20687
 FT CHAIN 20687 20727
 FT CHAIN 20727 20767
 FT CHAIN 20767 20807
 FT CHAIN 20807 20847
 FT CHAIN 20847 20887
 FT CHAIN 20887 20927
 FT CHAIN 20927 20967
 FT CHAIN 20967 21007
 FT CHAIN 21007 21047
 FT CHAIN 21047 21087
 FT CHAIN 21087 21127
 FT CHAIN 21127 21167
 FT CHAIN 21167 21207
 FT CHAIN 21207 21247
 FT CHAIN 21247 21287
 FT CHAIN 21287 21327
 FT CHAIN 21327 21367
 FT CHAIN 21367 21407
 FT CHAIN 21407 21447
 FT CHAIN 21447 21487
 FT CHAIN 21487 21527
 FT CHAIN 21527 21567
 FT CHAIN 21567 21607
 FT CHAIN 21607 21647
 FT CHAIN 21647 21687
 FT CHAIN 21687 21727
 FT CHAIN 21727 21767
 FT CHAIN 21767 21807
 FT CHAIN 21807 21847
 FT CHAIN 21847 21887
 FT CHAIN 21887 21927
 FT CHAIN 21927 21967
 FT CHAIN 21967 22007
 FT CHAIN 22007 22047
 FT CHAIN 22047 22087
 FT CHAIN 22087 22127
 FT CHAIN 22127 22167
 FT CHAIN 22167 22207
 FT CHAIN 22207 22247
 FT CHAIN 22247 22287
 FT CHAIN 22287 22327
 FT CHAIN 22327 22367
 FT CHAIN 22367 22407
 FT CHAIN 22407 22447
 FT CHAIN 22447 22487
 FT CHAIN 22487 22527
 FT CHAIN 22527 22567
 FT CHAIN 22567 22607
 FT CHAIN 22607 22647
 FT CHAIN 22647 22687
 FT CHAIN 22687 22727
 FT CHAIN 22727 22767
 FT CHAIN 22767 22807
 FT CHAIN 22807 22847
 FT CHAIN 22847 22887
 FT CHAIN 22887 22927
 FT CHAIN 22927 22967
 FT CHAIN 22967 23007
 FT CHAIN 23007 23047
 FT CHAIN 23047 23087
 FT CHAIN 23087 23127
 FT CHAIN 23127 23167
 FT CHAIN 23167 23207
 FT CHAIN 23207 23247
 FT CHAIN 23247 23287
 FT CHAIN 23287 23327
 FT CHAIN 23327 23367
 FT CHAIN 23367 23407
 FT CHAIN 23407 23447
 FT CHAIN 23447 23487
 FT CHAIN 23487 23527
 FT CHAIN 23527 23567
 FT CHAIN 23567 23607
 FT CHAIN 23607 23647
 FT CHAIN 23647 23687
 FT CHAIN 23687 23727
 FT CHAIN 23727 23767
 FT CHAIN 23767 23807
 FT CHAIN 23807 23847
 FT CHAIN 23847 23887
 FT CHAIN 23887 23927
 FT CHAIN 23927 23967
 FT CHAIN 23967 24007
 FT CHAIN 24007 24047
 FT CHAIN 24047 24087
 FT CHAIN 24087 24127
 FT CHAIN 24127 24167
 FT CHAIN 24167 24207
 FT CHAIN 24207 24247
 FT CHAIN 24247 24287
 FT CHAIN 24287 24327
 FT CHAIN 24327 24367
 FT CHAIN 24367 24407
 FT CHAIN 24407 24447
 FT CHAIN 24447 24487
 FT CHAIN 24487 24527
 FT CHAIN 24527 24567
 FT CHAIN 24567 24607
 FT CHAIN 24607 24647
 FT CHAIN 24647 24687
 FT CHAIN 24687 24727
 FT CHAIN 24727 24767
 FT CHAIN 24767 24807
 FT CHAIN 24807 24847
 FT CHAIN 24847 24887
 FT CHAIN 24887 24927
 FT CHAIN 24927 24967
 FT CHAIN 24967 25007
 FT CHAIN 25007 25047
 FT CHAIN 25047 25087
 FT CHAIN 25087 25127
 FT CHAIN 25127 25167
 FT CHAIN 25167 25207
 FT CHAIN 25207 25247
 FT CHAIN 25247 25287
 FT CHAIN 25287 25327
 FT CHAIN 25327 25367
 FT CHAIN 25367 25407
 FT CHAIN 25407 25447
 FT CHAIN 25447 25487
 FT CHAIN 25487 25527
 FT CHAIN 25527 25567
 FT CHAIN 25567 25607
 FT CHAIN 25607 25647
 FT CHAIN 25647 25687
 FT CHAIN 25687 25727
 FT CHAIN 25727 25767
 FT CHAIN 25767 25807
 FT CHAIN 25807 25847
 FT CHAIN 25847 25887
 FT CHAIN 25887 25927
 FT CHAIN 25927 25967
 FT CHAIN 25967 26007
 FT CHAIN 26007 26047
 FT CHAIN 26047 26087
 FT CHAIN 26087 26127
 FT CHAIN 26127 26167
 FT CHAIN 26167 26207
 FT CHAIN 26207 26247
 FT CHAIN 26247 26287
 FT CHAIN 26287 26327
 FT CHAIN 26327 26367
 FT CHAIN 26367 26407
 FT CHAIN 26407 26447
 FT CHAIN 26447 26487
 FT CHAIN 26487 26527
 FT CHAIN 26527 26567
 FT CHAIN 26567 26607
 FT CHAIN 26607 26647

[illegible]

| FT | DISULEFID | 1484 | 1497 | BY SIMILARITY. |
|----|-----------|----------|------------|-------------------------------------|
| FT | DISULEFID | 1616 | 1627 | BY SIMILARITY. |
| FT | DISULEFID | 1622 | 1636 | BY SIMILARITY. |
| FT | DISULEFID | 1638 | 1651 | BY SIMILARITY. |
| FT | DISULEFID | 1657 | 1672 | BY SIMILARITY. |
| FT | DISULEFID | 1667 | 1681 | BY SIMILARITY. |
| FT | DISULEFID | 1683 | 1696 | BY SIMILARITY. |
| FT | CARBOHYD | 339 | 339 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 370 | 370 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 416 | 416 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 612 | 612 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 1042 | 1042 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 1242 | 1242 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 1357 | 1357 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| SO | SEQUENCE | 1712 AA: | 186598 MM: | 650BCDEA691FDJ34 CRC64; |

| | | | | |
|-----------------------|------------------|--------------------|------------|--------------|
| Query Match | 11.2%; | Score 327.5; | DB 1; | Length 1712; |
| Best Local Similarity | 28.8%; | Pred. No. 2.7e-14; | | |
| Matches 106; | Conservative 44; | Mismatches 133; | Indels 85; | Gaps 25 |

| | | | | | | |
|----|------|--|-----------------------------------|----------------|---------|------|
| OY | 177 | PLAVEPGAAAAA | SVITTYGTFPAARGADPQALRPGS---- | SAANVRLGIOLM-- | CTAPRGA | 230 |
| Db | 831 | PVVVEKTSPPVPEVA----- | -----PBGSTSSASOYATAPQVETIECTVNPDI | | | 875 |
| OY | 231 | VQGHMARBARPAMDCSVENG----- | -----GCEHA----- | CNAIGAPRCOP | | 268 |
| Db | 876 | CGAGHCINLPVRYATCIEBEGYKFPSEQRKICID | DEDAQAHNLCSGORCENTBGSJLICP | | | 935 |
| OY | 269 | AGAAIQADGRSCTASATQSCNDLC-ENFECVPNDQDGSYC- | MCETGRILANDQRCEDV | | | 326 |
| Db | 936 | AGFIASBEGSMCIDVDELCIRPDVCRORCI-- | NTGAAPRCCECDSGYMSRGRH-CEDII | | | 991 |
| OY | 327 | DDCILEPSPCC-ORCVMTQGGFEC-PCYPUYDLVDECEVPYDPCRRAN- | CEYQSCPLN | | | 382 |
| Db | 992 | DEC-LPSTCPEECVSPQSTOCVPCSTEGSRGIMQCID-VDECLQPKVCTNGSCTIN- | | | | 1048 |
| OY | 383 | QTSYLICVAEGFAPIRPH-----RCOM--PCNQTACPADDPNTOAS-- | CESEGEYIL | | | 433 |
| Db | 1049 | EGSYIMCHKUYSTPRHRCODIDECQGNLCMNGQC----- | KNTDGSFRCTGGQGYQL | | | 1103 |
| OY | 434 | -----DDGFICTDIDECNNGGFCG-VCVCHNLPGFTEICGPD-- | SALVNHIG--TDC--D | | | 481 |
| Db | 1104 | SAAKDO--CEDIDCEHRLHCSHGOCRTNTEGSFOCLCNOGYRASVLDGHCEDINECTLED | | | | 1160 |
| OY | 482 | SGKVDGPD | 489 | | | |
| Db | 1161 | SSVCGGD | 1168 | | | |

Search completed: May 8, 2002, 12:45:14
Job time: 350 sec

Wed May 8 16:19:25 2002

us-09-509-994-1.rapm

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: May 8, 2002, 12:37:49 ; Search time 206.9 Seconds

(without alignments)
876.013 Million cell updates/sec

Title: US-09-509-994-1

Sequence: 1 MLGVLVGLALAGLGFPPAP.....PSPPGSGTLPPAVGLVHSG 516

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3516493 seqs, 351254056 residues

Total number of hits satisfying chosen parameters: 3516493

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main: *
1: /cgn2_6/ptodata/2/paa/PCRUS_COMB.pep: *
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep: *
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep: *
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep: *
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep: *
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep: *
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep: *
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep: *
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep: *
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep: *
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep: *
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep: *
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep: *
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep: *
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep: *
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep: *
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep: *
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep: *
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep: *
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep: *
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep: *
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep: *
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep: *
24: /cgn2_6/ptodata/2/paa/US060_COMB.pep: *
25: /cgn2_6/ptodata/2/paa/US061_COMB.pep: *
26: /cgn2_6/ptodata/2/paa/US10_MERGED_COMB.pep: *
27: /cgn2_6/ptodata/2/paa/US09_MERGED_COMB.pep: *
28: /cgn2_6/ptodata/2/paa/US09_MERGED_COMB.pep: *
29: /cgn2_6/ptodata/2/paa/US08_MERGED_COMB.pep: *
30: /cgn2_6/ptodata/2/paa/US07_MERGED_COMB.pep: *
31: /cgn2_6/ptodata/2/paa/US06_MERGED_COMB.pep: *
32: /cgn2_6/ptodata/2/paa/PCCT_MERGED_COMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------------------------|
| 1 | 2916 | 100.0 | 516 | 27 | US-09-509-994-1 Sequence 1, Appl1 |

ALIGNMENTS

RESULT 1
US-09-509-994-1
Sequence 1, Application US/09509994
GENERAL INFORMATION:
APPLICANT: YOKOZAWA, AKIRA
APPLICANT: YOTI, MASAKI
APPLICANT: MURATA, TOMOYO
APPLICANT: TSURUTA, KAZUHIRO
APPLICANT: SHIMIZU, HIROMOMO
TITLE OF INVENTION: METHOD FOR KEEPING THE QUALITY OF AQUEOUS PARENTERAL
FILE REFERENCE: KP-8753
CURRENT APPLICATION NUMBER: US/09/509,994
CURRENT FILING DATE: 2000-05-08
PRIOR APPLICATION NUMBER: PCT/JP98/04609
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: JP 9-281659
PRIOR FILING DATE: 1997-10-15
PRIOR APPLICATION NUMBER: JP 9-308523
PRIOR FILING DATE: 1997-11-11
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 516

| 2 | 2916 | 100.0 | 575 | 3 | US-07-689-936A-7 | Sequence 7, Appl1 |
|----|------|-------|-----|----|-------------------|--------------------|
| 3 | 2916 | 100.0 | 575 | 22 | US-09-880-464-2 | Sequence 2, Appl1 |
| 4 | 2916 | 100.0 | 575 | 23 | US-09-938-405-2 | Sequence 2, Appl1 |
| 5 | 2916 | 99.9 | 516 | 27 | US-09-509-994-2 | Sequence 2, Appl1 |
| 6 | 2912 | 99.9 | 575 | 3 | US-07-730-975-2 | Sequence 2, Appl1 |
| 7 | 2912 | 99.9 | 575 | 3 | US-07-948-759-4 | Sequence 4, Appl1 |
| 8 | 2912 | 99.9 | 575 | 23 | US-09-948-933-228 | Sequence 228, App |
| 9 | 2912 | 99.9 | 682 | 17 | US-09-948-933-318 | Sequence 318, App |
| 10 | 2824 | 96.8 | 497 | 17 | US-09-331-793-4 | Sequence 4, Appl1 |
| 11 | 2770 | 95.0 | 494 | 3 | US-07-796-336A-14 | Sequence 16, Appl |
| 12 | 2768 | 94.9 | 494 | 3 | US-07-796-336A-16 | Sequence 16, Appl |
| 13 | 2746 | 94.2 | 486 | 6 | US-08-212-384-4 | Sequence 4, Appl1 |
| 14 | 2690 | 92.2 | 475 | 3 | US-07-835-436C-1 | Sequence 1, Appl1 |
| 15 | 2687 | 92.1 | 475 | 3 | US-07-835-436C-1 | Sequence 1, Appl1 |
| 16 | 2687 | 92.1 | 475 | 3 | US-07-835-436A-1 | Sequence 1, Appl1 |
| 17 | 2686 | 92.1 | 475 | 3 | US-07-835-436A-2 | Sequence 2, Appl1 |
| 18 | 2686 | 92.1 | 475 | 3 | US-07-835-436A-2 | Sequence 2, Appl1 |
| 19 | 2686 | 92.1 | 475 | 3 | US-07-835-436C-2 | Sequence 2, Appl1 |
| 20 | 2684 | 92.0 | 476 | 3 | US-07-796-336A-1 | Sequence 1, Appl1 |
| 21 | 2682 | 92.0 | 476 | 3 | US-07-796-336A-17 | Sequence 17, Appl1 |
| 22 | 2665 | 91.4 | 468 | 6 | US-08-212-384-1 | Sequence 1, Appl1 |
| 23 | 2630 | 90.2 | 462 | 3 | US-07-897-194A-1 | Sequence 1, Appl1 |
| 24 | 2630 | 90.2 | 462 | 3 | US-07-897-194A-1 | Sequence 1, Appl1 |
| 25 | 2619 | 89.8 | 462 | 3 | US-07-897-194-3 | Sequence 3, Appl1 |
| 26 | 2615 | 89.7 | 462 | 3 | US-07-897-194-4 | Sequence 4, Appl1 |
| 27 | 2615 | 89.7 | 462 | 3 | US-07-897-194-5 | Sequence 5, Appl1 |
| 28 | 2615 | 89.7 | 462 | 3 | US-07-897-194-6 | Sequence 6, Appl1 |
| 29 | 2592 | 88.9 | 456 | 3 | US-07-835-436C-3 | Sequence 3, Appl1 |
| 30 | 2592 | 88.9 | 456 | 3 | US-07-835-436C-3 | Sequence 3, Appl1 |
| 31 | 2592 | 88.9 | 456 | 3 | US-07-835-436C-3 | Sequence 3, Appl1 |
| 32 | 2592 | 88.9 | 456 | 3 | US-07-835-436C-3 | Sequence 3, Appl1 |
| 33 | 2592 | 88.9 | 456 | 3 | US-07-835-436C-4 | Sequence 4, Appl1 |
| 34 | 2592 | 88.9 | 456 | 3 | US-07-835-436C-4 | Sequence 4, Appl1 |
| 35 | 2592 | 88.9 | 456 | 3 | US-07-835-436C-4 | Sequence 4, Appl1 |
| 36 | 2591 | 88.9 | 462 | 3 | US-07-897-194-10 | Sequence 10, Appl1 |
| 37 | 2576 | 88.3 | 462 | 3 | US-07-897-194-12 | Sequence 12, Appl1 |
| 38 | 2571 | 88.2 | 462 | 3 | US-07-897-194-16 | Sequence 16, Appl1 |
| 39 | 2571 | 88.2 | 462 | 3 | US-07-897-194-18 | Sequence 18, Appl1 |
| 40 | 2566 | 88.0 | 462 | 3 | US-07-897-194-14 | Sequence 14, Appl1 |
| 41 | 2561 | 87.8 | 462 | 3 | US-07-897-194-13 | Sequence 13, Appl1 |
| 42 | 2558 | 87.7 | 462 | 3 | US-07-897-194-12 | Sequence 12, Appl1 |
| 43 | 2543 | 87.2 | 446 | 3 | US-07-835-436-5 | Sequence 5, Appl1 |
| 44 | 2543 | 87.2 | 446 | 3 | US-07-835-436A-5 | Sequence 5, Appl1 |
| 45 | 2543 | 87.2 | 446 | 3 | US-07-835-436A-5 | Sequence 5, Appl1 |

```

; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Partial amino acid sequences of a human
; OTHER INFORMATION: thrombomodulin
US-09-509-994-1

```

```

Query Match          100.0%; Score 2916; DB 27; Length 516;
Best Local Similarity 100.0%; Pred. No. 4e-193;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVYVIGALALAGLGFPAAPAPQPGSGSQVEHDCFALYPPATFLNASQICDGLRGLM 60
DB 1 MGVYVIGALALAGLGFPAAPAPQPGSGSQVEHDCFALYPPATFLNASQICDGLRGLM 60
QY 61 TVRSSVAADVISILLNGDGVGRRRLMIGLQLPFGCGDPKRLGRLGQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISILLNGDGVGRRRLMIGLQLPFGCGDPKRLGRLGQWVTGDNNTSYS 120
QY 121 RMARLDLNGAPLGGPLCVAVSAEAATVPSEPIWEBOQCEVKAADGFLCEHFHPATCRPLAV 180
DB 121 RMARLDLNGAPLGGPLCVAVSAEAATVPSEPIWEBOQCEVKAADGFLCEHFHPATCRPLAV 180
QY 181 EPGAAAANVSTYGTTPFAARGADFOALPVGSSAANAAPLGLQIMCTAPGAVQGHAREAP 240
DB 181 EPGAAAANVSTYGTTPFAARGADFOALPVGSSAANAAPLGLQIMCTAPGAVQGHAREAP 240
QY 241 GAMDCSVENGSCHEACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGSCHEACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCEGTGYRLAADQHRCEVDVDCILPSPQPCRCVNTQGGFECHCYPNTDLYDG 360
DB 301 DPGSYSCMCEGTGYRLAADQHRCEVDVDCILPSPQPCRCVNTQGGFECHCYPNTDLYDG 360
QY 361 BCEVEVDQCFRANCYOCOPLNQSTYLCVABGAPRTPHEBRCOMFONQACPADCDPN 420
DB 361 BCEVEVDQCFRANCYOCOPLNQSTYLCVABGAPRTPHEBRCOMFONQACPADCDPN 420
QY 421 TQASCECEGYLLDDGFLCTDIDECENGFCSGVCHNLPGTFEFCIGPDSALVHIGTDC 480
DB 421 TQASCECEGYLLDDGFLCTDIDECENGFCSGVCHNLPGTFEFCIGPDSALVHIGTDC 480
QY 481 DSGKVDGSDSGSEPPSPPTPGSTLTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSEPPSPPTPGSTLTPPAVGLVHSG 516

```

RESULT 2

```

; Sequence 7, Application US/07689936A
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Shuji
; APPLICANT: Goml, Komakazu
; APPLICANT: Ogawa, Kohel
; TITLE OF INVENTION: An Isolated Physiologically Active
; TITLE OF INVENTION: Human Thrombomodulin Polypeptide
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington Street P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/689,936A

```

```

; FILING DATE: 19910521
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP90/01234
; FILING DATE: 25-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, Leonard R.
; REGISTRATION NUMBER: 30,330
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-0369
; TELEX: 248345
; INFORMATION FOR SEQ. ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..575
; OTHER INFORMATION: /label= protein
; OTHER INFORMATION: /note= "thrombomodulin preproprotein"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..18
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "signal peptide"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 517..575
; OTHER INFORMATION: /label= domain
; OTHER INFORMATION: /note= "membrane-binding and cytoplasmic domains"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 366..480
; OTHER INFORMATION: /label= domain
; OTHER INFORMATION: /note= "enzyme active site and chondroitin or
; OTHER INFORMATION: chondroitin-sulfate attachment region"
US-07-689-936A-7

```

```

Query Match          100.0%; Score 2916; DB 3; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.5e-193;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MGVYVIGALALAGLGFPAAPAPQPGSGSQVEHDCFALYPPATFLNASQICDGLRGLM 60
DB 1 MGVYVIGALALAGLGFPAAPAPQPGSGSQVEHDCFALYPPATFLNASQICDGLRGLM 60
QY 61 TVRSSVAADVISILLNGDGVGRRRLMIGLQLPFGCGDPKRLGRLGQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISILLNGDGVGRRRLMIGLQLPFGCGDPKRLGRLGQWVTGDNNTSYS 120
QY 121 RMARLDLNGAPLGGPLCVAVSAEAATVPSEPIWEBOQCEVKAADGFLCEHFHPATCRPLAV 180
DB 121 RMARLDLNGAPLGGPLCVAVSAEAATVPSEPIWEBOQCEVKAADGFLCEHFHPATCRPLAV 180
QY 181 EPGAAAANVSTYGTTPFAARGADFOALPVGSSAANAAPLGLQIMCTAPGAVQGHAREAP 240
DB 181 EPGAAAANVSTYGTTPFAARGADFOALPVGSSAANAAPLGLQIMCTAPGAVQGHAREAP 240
QY 241 GAMDCSVENGSCHEACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGSCHEACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCEGTGYRLAADQHRCEVDVDCILPSPQPCRCVNTQGGFECHCYPNTDLYDG 360
DB 301 DPGSYSCMCEGTGYRLAADQHRCEVDVDCILPSPQPCRCVNTQGGFECHCYPNTDLYDG 360

```

Wed May 8 16:19:25 2002

us-09-509-994-1.1.ram

Page 3

OY 361 ECVEPVDPCEFRANCEYOCQPLNOTSYLVCACGAFAPRPHRHRCOMFCNOTACPADCDPN 420
DB 361 ECVEPVDPCEFRANCEYOCQPLNOTSYLVCACGAFAPRPHRHRCOMFCNOTACPADCDPN 420
OY 421 TOASCECEPEGYIIDDGFICTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVRIHIGTDC 480
DB 421 TOASCECEPEGYIIDDGFICTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVRIHIGTDC 480
OY 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
RESULT 3
US-09-880-464-2 099880484
; Sequence 2, Application US/09880464
; GENERAL INFORMATION:
; APPLICANT: Light, David
; APPLICANT: Morser, Michael J
; APPLICANT: Nagashima, Marko
; TITLE OF INVENTION: Thrombomodulin Analogs for Pharmaceutical Use
; FILE REFERENCE: Improved Thrombomodulin Analogs
; CURRENT APPLICATION NUMBER: US/09/880,464
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/213,678
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-464-2

Query Match 100.0%; Score 2916; DB 22; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.5e-193;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MGLVLYGALALAGLGFPAPEPQSGSCVEHDCFALYGPATFLNASQICDGLRGLHM 60
DB 1 MGLVLYGALALAGLGFPAPEPQSGSCVEHDCFALYGPATFLNASQICDGLRGLHM 60
OY 61 TVRSSVADYISLLNDGGVGRRLMIGLQLPFGCGDPRKGLRFGOMTGDNNNTSYS 120
DB 61 TVRSSVADYISLLNDGGVGRRLMIGLQLPFGCGDPRKGLRFGOMTGDNNNTSYS 120
OY 121 RMARLDNGAPLGLPCLVAVSAEATVPSEPIWEBOOCEVKADGFLCEHFHPRATCRPLAV 180
DB 121 RMARLDNGAPLGLPCLVAVSAEATVPSEPIWEBOOCEVKADGFLCEHFHPRATCRPLAV 180
OY 181 EPGAAAAVSTYGTTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPPAVAGHMAAREAP 240
DB 181 EPGAAAAVSTYGTTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPPAVAGHMAAREAP 240
OY 241 GAMDCEVNGGCEHACNAIPGARCCOPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCEVNGGCEHACNAIPGARCCOPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
OY 301 DPGSYSCMCEYRLAADHRCEDVDCTLEPSPCQRCVNTGGFECCHCYPNYDLVDG 360
DB 301 DPGSYSCMCEYRLAADHRCEDVDCTLEPSPCQRCVNTGGFECCHCYPNYDLVDG 360
OY 361 ECVEPVDPCEFRANCEYOCQPLNOTSYLVCACGAFAPRPHRHRCOMFCNOTACPADCDPN 420
DB 361 ECVEPVDPCEFRANCEYOCQPLNOTSYLVCACGAFAPRPHRHRCOMFCNOTACPADCDPN 420
OY 421 TOASCECEPEGYIIDDGFICTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVRIHIGTDC 480
DB 421 TOASCECEPEGYIIDDGFICTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVRIHIGTDC 480
OY 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516

DB 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
RESULT 4
US-09-938-405-2
; Sequence 2, Application US/09938405
; GENERAL INFORMATION:
; APPLICANT: Pestoff, Barry W.
; APPLICANT: Morser, Michael J.
; TITLE OF INVENTION: Thrombomodulin Analogs for Use in Recovery of Spinal Cord Inju
; FILE REFERENCE: 51960AUSM1
; CURRENT APPLICATION NUMBER: US/09/938,405
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 60/229,714
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 2
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-938-405-2

Query Match 100.0%; Score 2916; DB 23; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.5e-193;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MGLVLYGALALAGLGFPAPEPQSGSCVEHDCFALYGPATFLNASQICDGLRGLHM 60
DB 1 MGLVLYGALALAGLGFPAPEPQSGSCVEHDCFALYGPATFLNASQICDGLRGLHM 60
OY 61 TVRSSVADYISLLNDGGVGRRLMIGLQLPFGCGDPRKGLRFGOMTGDNNNTSYS 120
DB 61 TVRSSVADYISLLNDGGVGRRLMIGLQLPFGCGDPRKGLRFGOMTGDNNNTSYS 120
OY 121 RMARLDNGAPLGLPCLVAVSAEATVPSEPIWEBOOCEVKADGFLCEHFHPRATCRPLAV 180
DB 121 RMARLDNGAPLGLPCLVAVSAEATVPSEPIWEBOOCEVKADGFLCEHFHPRATCRPLAV 180
OY 181 EPGAAAAVSTYGTTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPPAVAGHMAAREAP 240
DB 181 EPGAAAAVSTYGTTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPPAVAGHMAAREAP 240
OY 241 GAMDCEVNGGCEHACNAIPGARCCOPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCEVNGGCEHACNAIPGARCCOPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
OY 301 DPGSYSCMCEYRLAADHRCEDVDCTLEPSPCQRCVNTGGFECCHCYPNYDLVDG 360
DB 301 DPGSYSCMCEYRLAADHRCEDVDCTLEPSPCQRCVNTGGFECCHCYPNYDLVDG 360
OY 361 ECVEPVDPCEFRANCEYOCQPLNOTSYLVCACGAFAPRPHRHRCOMFCNOTACPADCDPN 420
DB 361 ECVEPVDPCEFRANCEYOCQPLNOTSYLVCACGAFAPRPHRHRCOMFCNOTACPADCDPN 420
OY 421 TOASCECEPEGYIIDDGFICTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVRIHIGTDC 480
DB 421 TOASCECEPEGYIIDDGFICTDIDCEGNGFCSCVCHNLPTGTEFCICGPDALVRIHIGTDC 480
OY 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
RESULT 5
US-09-509-994-2
; Sequence 2, Application US/09509994
; GENERAL INFORMATION:
; APPLICANT: YUI, MASAKI
; APPLICANT: YOKOZAKA, AKIRA
; APPLICANT: MORATA, TOMOYO
; APPLICANT: TSURUTA, KAZUHISA

```

; APPLICANT: SHIMIZU, HIROKOMO
; TITLE OF INVENTION: METHOD FOR KEEPING THE QUALITY OF AQUEOUS PARENTERAL
; FILE OF INVENTION: SOLUTION OF THROMBOMODULIN IN STORAGE AND DISTRIBUTION
; FILE REFERENCE: KP-8753
; CURRENT APPLICATION NUMBER: US/09/509,994
; CURRENT FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: PCT/JP98/04609
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: JP 9-281659
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: JP 9-308523
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Partial amino acid sequences of a human
; OTHER INFORMATION: Thrombomodulin
US-09-509-994-2

```

```

Query Match          99.98%  Score 2912;  DB 27;  Length 516;
Best Local Similarity 99.88%  Pred. No. 7.5e-193;
Matches 515;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

```

```

QY 1 MLCVIVGALALAGLGFPAPEPOPGSGQVEHDCFALYPGATFLNLSQICDGLRGLHM 60
DB 1 MLCVIVGALALAGLGFPAPEPOPGSGQVEHDCFALYPGATFLNLSQICDGLRGLHM 60
QY 61 TVSSVAADYISLLNGDGVGRRRLMIGQLPPGCGDPKRLGLPGFOWVTGDNNTSYS 120
DB 61 TVSSVAADYISLLNGDGVGRRRLMIGQLPPGCGDPKRLGLPGFOWVTGDNNTSYS 120
QY 121 RMARLDLNGAPLGLCPVAVSAEATVPSEPIWEEOCEYKADGFLCEPHFPATCRPLAV 180
DB 121 RMARLDLNGAPLGLCPVAVSAEATVPSEPIWEEOCEYKADGFLCEPHFPATCRPLAV 180
QY 121 RMARLDLNGAPLGLCPVAVSAEATVPSEPIWEEOCEYKADGFLCEPHFPATCRPLAV 180
DB 121 RMARLDLNGAPLGLCPVAVSAEATVPSEPIWEEOCEYKADGFLCEPHFPATCRPLAV 180
QY 181 EPGAANAASITTYGTPPAAGADFOALPVGSSAANAVALGLQIMCTAPGAVOGHMAREAP 240
DB 181 EPGAANAASITTYGTPPAAGADFOALPVGSSAANAVALGLQIMCTAPGAVOGHMAREAP 240
QY 241 GAMDCSVENGCEHACNAITGAPRCOCFAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGCEHACNAITGAPRCOCFAGALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILFSPCPCRCVNTQGGFECCHYPNYDLYDG 360
DB 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILFSPCPCRCVNTQGGFECCHYPNYDLYDG 360
QY 361 ECEVPDPCFRANCEYQCQPLNQTSTLYCAEGFAPRIPEPHRCQMFNCQTACPADCDPN 420
DB 361 ECEVPDPCFRANCEYQCQPLNQTSTLYCAEGFAPRIPEPHRCQMFNCQTACPADCDPN 420
QY 421 TOASCCEPEGYIIDDGFTICDIDECENGCGFSGVCHNLPTGTFECICGPPSALVRIHGTDC 480
DB 421 TOASCCEPEGYIIDDGFTICDIDECENGCGFSGVCHNLPTGTFECICGPPSALVRIHGTDC 480
QY 481 DSGKVDGSDSGSEPPSPPTPGSTLTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSEPPSPPTPGSTLTPPAVGLVHSG 516

```

```

RESULT 6
US-07-730-975-2
; Sequence 2, Application US/07730975
; GENERAL INFORMATION:
; APPLICANT: MORSE, MICHAEL J.
; APPLICANT: GLASER, CHARLES
; TITLE OF INVENTION: SOLUBLE ANALOGS OF THROMBOMODULIN
; NUMBER OF SEQUENCES: 27

```

```

CORRESPONDENCE ADDRESS:
; ADDRESS: BERLEX BIOSCIENCES
; STREET: 1501 HARBOR BAY PARKWAY
; CITY: ALAMEDA
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94501
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/730,975
; FILING DATE: 19910729
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: CHING, EDWIN P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: A-0155A US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 266-7476
; TELEFAX: (415) 266-7400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-730-975-2

```

```

Query Match          99.98%  Score 2912;  DB 3;  Length 575;
Best Local Similarity 99.88%  Pred. No. 8.5e-193;
Matches 515;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

```

```

QY 1 MLCVIVGALALAGLGFPAPEPOPGSGQVEHDCFALYPGATFLNLSQICDGLRGLHM 60
DB 1 MLCVIVGALALAGLGFPAPEPOPGSGQVEHDCFALYPGATFLNLSQICDGLRGLHM 60
QY 61 TVSSVAADYISLLNGDGVGRRRLMIGQLPPGCGDPKRLGLPGFOWVTGDNNTSYS 120
DB 61 TVSSVAADYISLLNGDGVGRRRLMIGQLPPGCGDPKRLGLPGFOWVTGDNNTSYS 120
QY 121 RMARLDLNGAPLGLCPVAVSAEATVPSEPIWEEOCEYKADGFLCEPHFPATCRPLAV 180
DB 121 RMARLDLNGAPLGLCPVAVSAEATVPSEPIWEEOCEYKADGFLCEPHFPATCRPLAV 180
QY 121 RMARLDLNGAPLGLCPVAVSAEATVPSEPIWEEOCEYKADGFLCEPHFPATCRPLAV 180
DB 121 RMARLDLNGAPLGLCPVAVSAEATVPSEPIWEEOCEYKADGFLCEPHFPATCRPLAV 180
QY 181 EPGAANAASITTYGTPPAAGADFOALPVGSSAANAVALGLQIMCTAPGAVOGHMAREAP 240
DB 181 EPGAANAASITTYGTPPAAGADFOALPVGSSAANAVALGLQIMCTAPGAVOGHMAREAP 240
QY 241 GAMDCSVENGCEHACNAITGAPRCOCFAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGCEHACNAITGAPRCOCFAGALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILFSPCPCRCVNTQGGFECCHYPNYDLYDG 360
DB 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILFSPCPCRCVNTQGGFECCHYPNYDLYDG 360
QY 361 ECEVPDPCFRANCEYQCQPLNQTSTLYCAEGFAPRIPEPHRCQMFNCQTACPADCDPN 420
DB 361 ECEVPDPCFRANCEYQCQPLNQTSTLYCAEGFAPRIPEPHRCQMFNCQTACPADCDPN 420
QY 421 TOASCCEPEGYIIDDGFTICDIDECENGCGFSGVCHNLPTGTFECICGPPSALVRIHGTDC 480
DB 421 TOASCCEPEGYIIDDGFTICDIDECENGCGFSGVCHNLPTGTFECICGPPSALVRIHGTDC 480
QY 481 DSGKVDGSDSGSEPPSPPTPGSTLTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSEPPSPPTPGSTLTPPAVGLVHSG 516

```

RESULT 7
US-07-949-759-4
; Sequence 4, Application US/07949759
; GENERAL INFORMATION:
; APPLICANT: Beeler, David L.
; APPLICANT: Jackman, Robert W
; APPLICANT: Rosenberg, Robert B
; APPLICANT: Fritze, Linda
; APPLICANT: Soff, Gerald
; TITLE OF INVENTION: Thrombomodulin of Mammalian Origin
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Milltia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/949,759
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/675,307
; FILING DATE: 26-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MT-4545A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-949-759-4

Query Match 99.98; Score 2912; DB 3; Length 575;
Best Local Similarity 99.88; Pred. No. 8.5e-193;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 1 MGVLVIGALALAGLGFPAAPAPQPGSGQCVHEHDCFALYGPATFLNASQICDGLRGHLM 60
Db 1 MGVLVIGALALAGLGFPAAPAPQPGSGQCVHEHDCFALYGPATFLNASQICDGLRGHLM 60
Qy 61 TYRSSVADVISTLLNGGCGVGRRLWIGLQLPFGCGDPRKLGPRGFOWTGDNNISYS 120
Db 61 TYRSSVADVISTLLNGGCGVGRRLWIGLQLPFGCGDPRKLGPRGFOWTGDNNISYS 120
Qy 121 RWARLDLNGAPLGCPLCAVAASAETVPSPIWEBOCEVKAADGFLCEFHFPATCRPLAV 180
Db 121 RWARLDLNGAPLGCPLCAVAASAETVPSPIWEBOCEVKAADGFLCEFHFPATCRPLAV 180
Qy 181 EFGAAAVASTITGTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHWAAREAP 240
Db 181 EFGAAAVASTITGTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHWAAREAP 240
Qy 241 GAMDCSVENGCGEHAACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAMDCSVENGCGEHAACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
Qy 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILPESPCPCRCVNTQSGFECCHYPNYDLVDG 360
Db 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILPESPCPCRCVNTQSGFECCHYPNYDLVDG 360

Qy 361 ECVEPVDPCEFRANCEYOCQPLNQTSTLYLCAEGFAPIPHEHRNRCMFCNQTACPADCPN 420
Db 361 ECVEPVDPCEFRANCEYOCQPLNQTSTLYLCAEGFAPIPHEHRNRCMFCNQTACPADCPN 420
Qy 421 TOASCECPREGIILDDGFICTDIDCENGGFCGVCCHNLPGTFECICGPDALVRIHGTDC 480
Db 421 TOASCECPREGIILDDGFICTDIDCENGGFCGVCCHNLPGTFECICGPDALVRIHGTDC 480
Qy 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAGLVHSG 516
Db 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAGLVHSG 516

RESULT 8
US-09-948-933-228
; Sequence 228, Application US/09948933
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH CARDIOVASCULAR DISEASE AND LIPID DISORDERS, METHODS OF
; DETECTION AND USES THEREOF
; FILE REFERENCE: C1000787
; CURRENT APPLICATION NUMBER: US/09/948,933
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,399
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 6404
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 228
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Human
US-09-948-933-228

Query Match 99.98; Score 2912; DB 23; Length 575;
Best Local Similarity 99.88; Pred. No. 8.5e-193;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MGVLVIGALALAGLGFPAAPAPQPGSGQCVHEHDCFALYGPATFLNASQICDGLRGHLM 60
Db 1 MGVLVIGALALAGLGFPAAPAPQPGSGQCVHEHDCFALYGPATFLNASQICDGLRGHLM 60
Qy 61 TYRSSVADVISTLLNGGCGVGRRLWIGLQLPFGCGDPRKLGPRGFOWTGDNNISYS 120
Db 61 TYRSSVADVISTLLNGGCGVGRRLWIGLQLPFGCGDPRKLGPRGFOWTGDNNISYS 120
Qy 121 RWARLDLNGAPLGCPLCAVAASAETVPSPIWEBOCEVKAADGFLCEFHFPATCRPLAV 180
Db 121 RWARLDLNGAPLGCPLCAVAASAETVPSPIWEBOCEVKAADGFLCEFHFPATCRPLAV 180
Qy 181 EFGAAAVASTITGTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHWAAREAP 240
Db 181 EFGAAAVASTITGTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHWAAREAP 240
Qy 241 GAMDCSVENGCGEHAACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAMDCSVENGCGEHAACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
Qy 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILPESPCPCRCVNTQSGFECCHYPNYDLVDG 360
Db 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILPESPCPCRCVNTQSGFECCHYPNYDLVDG 360
Qy 361 ECVEPVDPCEFRANCEYOCQPLNQTSTLYLCAEGFAPIPHEHRNRCMFCNQTACPADCPN 420
Db 361 ECVEPVDPCEFRANCEYOCQPLNQTSTLYLCAEGFAPIPHEHRNRCMFCNQTACPADCPN 420
Qy 421 TOASCECPREGIILDDGFICTDIDCENGGFCGVCCHNLPGTFECICGPDALVRIHGTDC 480
Db 421 TOASCECPREGIILDDGFICTDIDCENGGFCGVCCHNLPGTFECICGPDALVRIHGTDC 480
Qy 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAGLVHSG 516

Db 481 DSGKVDGDSGSGEPSPPTPGSTLTPPAVGLVHSG 516

|||||

RESULT 9

US-09-948-933-318

; Sequence 318, Application US/09948933

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH CARDIOVASCULAR DISEASE AND LIPID DISORDERS, METHODS OF

; TITLE OF INVENTION: DETECTION AND USES THEREOF

; FILE REFERENCE: CL000787

; CURRENT APPLICATION NUMBER: US/09/948,933

; CURRENT FILING DATE: 2001-09-10

; PRIOR APPLICATION NUMBER: 60/231,399

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 6404

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 318

; LENGTH: 682

; TYPE: PRT

; ORGANISM: Human

US-09-948-933-318

Query Match 99.9%; Score 2912; DB 23; Length 682;

Best Local Similarity 99.8%; Pred. No. 1e-192;

Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLYVIGALALAGLGPAPAEPOGSGQVEHDCFLALPGPATFLNASQICDGLRGLIM 60

Db 108 MGVLYVIGALALAGLGPAPAEPOGSGQVEHDCFLALPGPATFLNASQICDGLRGLIM 167

QY 61 TVRSSVADYISLLNGDGGVRRRLMIGLQPLPGCGDKRLGLPLGFOWTGDNNNTSYS 120

Db 168 TVRSSVADYISLLNGDGGVRRRLMIGLQPLPGCGDKRLGLPLGFOWTGDNNNTSYS 227

QY 121 RMARLDINGAPLCGLCAVAASAEATVPSEPIWEEOCEKADGFLCEHFHPATCRPLAV 180

Db 228 RMARLDINGAPLCGLCAVAASAEATVPSEPIWEEOCEKADGFLCEHFHPATCRPLAV 287

QY 181 EPGAAAASVITGTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREP 240

Db 288 EPGAAAASVITGTPFAAGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREP 347

QY 241 GAMDCEVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPP 300

Db 348 GAMDCEVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPP 407

QY 301 DQPGSYSCMCEGYRLAADHRCEDVDCTLEPSPQRCVNTQGGFECHECPNVDLVG 360

Db 408 DQPGSYSCMCEGYRLAADHRCEDVDCTLEPSPQRCVNTQGGFECHECPNVDLVG 467

QY 361 ECEVPVDFRANCEYOCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQOTACPADCDPN 420

Db 468 ECEVPVDFRANCEYOCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQOTACPADCDPN 527

QY 421 TQASCECPBEYIILDGFTCTDIDECENGFCGVCNHLPGTEFECICGPDALVRHIGTDC 480

Db 528 TQASCECPBEYIILDGFTCTDIDECENGFCGVCNHLPGTEFECICGPDALVRHIGTDC 587

QY 481 DSGKVDGDSGSGEPSPPTPGSTLTPPAVGLVHSG 516

Db 588 DSGKVDGDSGSGEPSPPTPGSTLTPPAVGLVHSG 623

RESULT 10

US-09-331-793-4

; Sequence 4, Application US/09331793

; GENERAL INFORMATION:

; APPLICANT: KURIGAMA, Shinichi

; APPLICANT: HASEGAMA, Takashi

; TITLE OF INVENTION: CELL MEMBRANE DIRECTED DRUGS

; FILE REFERENCE: 1110-253P

; CURRENT APPLICATION NUMBER: US/09/331,793

; CURRENT FILING DATE: 1999-06-25

; NUMBER OF SEQ ID NOS: 67

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4

; LENGTH: 497

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-331-793-4

Query Match 96.8%; Score 2824; DB 17; Length 497;

Best Local Similarity 100.0%; Pred. No. 8.8e-187;

Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 APAEPQSGSCVHDCFLALPGPATFLNASQICDGLRGLMTVRSSVADYISLLND 78

Db 1 APAEPQSGSCVHDCFLALPGPATFLNASQICDGLRGLMTVRSSVADYISLLND 60

QY 79 GGVRRRLMIGLQPLPGCGDKRLGLPLGRFOWTGDNNNTSYSRMARLDINGAPLCGLCY 138

Db 61 GGVRRRLMIGLQPLPGCGDKRLGLPLGRFOWTGDNNNTSYSRMARLDINGAPLCGLCY 120

QY 139 AVSAAEATVPSEPIWEEOCEKADGFLCEHFHPATCRPLAVEPGAAAASVITGTPFA 198

Db 121 AVSAAEATVPSEPIWEEOCEKADGFLCEHFHPATCRPLAVEPGAAAASVITGTPFA 180

QY 199 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREPAMDCSVENGCEHACNA 258

Db 181 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVOGHMAREPAMDCSVENGCEHACNA 240

QY 259 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPPQPSYSCMCEGYRLAA 318

Db 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPPQPSYSCMCEGYRLAA 300

QY 319 DQHRCEVDCTLEPSPQRCVNTQGGFECHECPNVDLVGCEVPVDFRANCEYOC 378

Db 301 DQHRCEVDCTLEPSPQRCVNTQGGFECHECPNVDLVGCEVPVDFRANCEYOC 360

QY 379 QPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQOTACPADCDPNQASCEPEGYIILDGFT 438

Db 361 QPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQOTACPADCDPNQASCEPEGYIILDGFT 420

QY 439 CTDIDECENGFCGVCNHLPGTEFECICGPDALVRHIGTDCSGKVDGDSGSGEPSP 498

Db 421 CTDIDECENGFCGVCNHLPGTEFECICGPDALVRHIGTDCSGKVDGDSGSGEPSP 480

QY 499 PTGSTLTPPAVGLVHSG 515

Db 481 PTGSTLTPPAVGLVHSG 497

RESULT 11

US-07-796-336A-14

; Sequence 14, Application US/07796336A

; GENERAL INFORMATION:

; APPLICANT: DOI, Takeshi

; APPLICANT: Iwasaki, Akio

; APPLICANT: Saino, Yushi

; APPLICANT: Kimura, Shigeru

; APPLICANT: Ohkuchi, Masao

; TITLE OF INVENTION: Thrombin-Binding Substance and Process

; TITLE OF INVENTION: For Preparing the Same

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLOM, SPIVAK, MCLELLAND, MAIER & NEUSTADT,

; STREET: P. C.

; CITY: Arlington

; STATE: Virginia

; ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/796.336A
FILING DATE: 19911122
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, Norman F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-066-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-796-336A-14

Query Match 95.0%; Score 2770; DB 3; Length 494;
Best Local Similarity 99.4%; Pred. No. 4.8e-183;
Matches 491; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVLYLGNALAGLGFPAAPAPPOPGSQCYEHDCFFALYPPAPFLNASQICDGLRSHLM 60
DB 1 MGVLYLGNALAGLGFPAAPAPPOPGSQCYEHDCFFALYPPAPFLNASQICDGLRSHLM 60
QY 61 TVRSSVAADVLSLLNGGGVGRRLWTGLQLPGCGDPKRLGRLGFOWVTGDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGGGVGRRLWTGLQLPGCGDPKRLGRLGFOWVTGDNNTSYS 120
QY 121 RPARLDLNGAPLCGPLCAVNSAETVSEPIWEBOCEYKADGFLCEFHPPATCRPLAV 180
DB 121 RPARLDLNGAPLCGPLCAVNSAETVSEPIWEBOCEYKADGFLCEFHPPATCRPLAV 180
QY 181 EFGAAAAAASITGTTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGMAREAP 240
DB 181 EFGAAAAAASITGTTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGMAREAP 240
QY 241 GAMDCSVENGCEHACNAPGAPRCQCPAGALADGHSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGCEHACNAPGAPRCQCPAGALADGHSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCEGTGRLADQHRCEEDVDDCIIEPSPCPCRCVNTGSGFECCHCYPNYDLVDG 360
DB 301 DPGSYSCMCEGTGRLADQHRCEEDVDDCIIEPSPCPCRCVNTGSGFECCHCYPNYDLVDG 360
QY 361 ECEVPDPCFRANCEYOCOPLNQTSYLCVCAEGFAPLIPHEPHRCOMFCNOTACPADCPN 420
DB 361 ECEVPDPCFRANCEYOCOPLNQTSYLCVCAEGFAPLIPHEPHRCOMFCNOTACPADCPN 420
QY 421 TQASCECPRGYLLDGGFICTDIDECENGSGVGHNLPGFECCIGGDSALVHHTGDC 480
DB 421 TQASCECPRGYLLDGGFICTDIDECENGSGVGHNLPGFECCIGGDSALVHHTGDC 480
QY 481 DSGKVDGDSGSGE 494
DB 481 DSGKVDGDSGSGE 494

RESULT 12
US-07-796-336A-16
Sequence 16, Application US/07796336A
GENERAL INFORMATION:
APPLICANT: Dol, Takeshi
APPLICANT: Iwasaki, Akio

APPLICANT: Saino, Yushi
APPLICANT: Kimura, Shigeru
APPLICANT: Ohkuchi, Masao
TITLE OF INVENTION: Thrombin-Binding Substance and Process
TITLE OF INVENTION: For Preparing the Same
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: P. C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/796.336A
FILING DATE: 19911122
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, Norman F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-066-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-796-336A-16

Query Match 94.9%; Score 2768; DB 3; Length 494;
Best Local Similarity 99.2%; Pred. No. 6.5e-183;
Matches 490; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVLYLGNALAGLGFPAAPAPPOPGSQCYEHDCFFALYPPAPFLNASQICDGLRSHLM 60
DB 1 MGVLYLGNALAGLGFPAAPAPPOPGSQCYEHDCFFALYPPAPFLNASQICDGLRSHLM 60
QY 61 TVRSSVAADVLSLLNGGGVGRRLWTGLQLPGCGDPKRLGRLGFOWVTGDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGGGVGRRLWTGLQLPGCGDPKRLGRLGFOWVTGDNNTSYS 120
QY 121 RPARLDLNGAPLCGPLCAVNSAETVSEPIWEBOCEYKADGFLCEFHPPATCRPLAV 180
DB 121 RPARLDLNGAPLCGPLCAVNSAETVSEPIWEBOCEYKADGFLCEFHPPATCRPLAV 180
QY 181 EFGAAAAAASITGTTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGMAREAP 240
DB 181 EFGAAAAAASITGTTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGMAREAP 240
QY 241 GAMDCSVENGCEHACNAPGAPRCQCPAGALADGHSCTASATQSCNDLCEHFCVNP 300
DB 241 GAMDCSVENGCEHACNAPGAPRCQCPAGALADGHSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCEGTGRLADQHRCEEDVDDCIIEPSPCPCRCVNTGSGFECCHCYPNYDLVDG 360
DB 301 DPGSYSCMCEGTGRLADQHRCEEDVDDCIIEPSPCPCRCVNTGSGFECCHCYPNYDLVDG 360
QY 361 ECEVPDPCFRANCEYOCOPLNQTSYLCVCAEGFAPLIPHEPHRCOMFCNOTACPADCPN 420
DB 361 ECEVPDPCFRANCEYOCOPLNQTSYLCVCAEGFAPLIPHEPHRCOMFCNOTACPADCPN 420
QY 421 TQASCECPRGYLLDGGFICTDIDECENGSGVGHNLPGFECCIGGDSALVHHTGDC 480
DB 421 TQASCECPRGYLLDGGFICTDIDECENGSGVGHNLPGFECCIGGDSALVHHTGDC 480

Db 421 TQASCEPEGYIIDDGFICTDIDECENGFGCSVCCHNLPGTFECTICGPDALVRIHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
Db 481 DSGKVDDEASGSGD 494

RESULT 13

Sequence 4, Application US/08212384
GENERAL INFORMATION:
APPLICANT: KIMURA, SHIGERU
APPLICANT: MIZOGUCHI, TOSHIMI
APPLICANT: OKUCHI, MASAO
APPLICANT: DOI, TAKESHI
TITLE OF INVENTION: PREPARATION OF A THROMBIN-BINDING
SUBSTANCE
TITLE OF INVENTION: SUBSTANCE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: OHLOM, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
City: Arlington
State: Virginia
Country: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,384
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,691
FILING DATE: 06-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Ohlom, Norman F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-063-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)412-3000
TELEFAX: (703)413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-212-384-4

Query Match 94.2%; Score 2746; DB 6; Length 486;
Best Local Similarity 99.8%; Pred. No. 2.1e-181;
Matches 485; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGVIVLGLALAGLIGFPAPAEPOPGSGSCVEHDFALYPGPATFLNASQICDGLRGHLM 60
Db 1 MGVIVLGLALAGLIGFPAPAEPOPGSGSCVEHDFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TYRSSVADVYSLILNDGSGVGRRLMTIGQLPPGCGDGRKRLGLPGFOWVTGDNNTSYS 120
Db 61 TYRSSVADVYSLILNDGSGVGRRLMTIGQLPPGCGDGRKRLGLPGFOWVTGDNNTSYS 120
QY 121 RWRARDLNGAPLGCPLCAVSAARATVPSEPIWEEOCEVKADGFLCEFHFPATGRPLAV 180
Db 121 RWRARDLNGAPLGCPLCAVSAARATVPSEPIWEEOCEVKADGFLCEFHFPATGRPLAV 180
QY 161 EPGAAAAAVSTYGTGTPFAARGADFOALPVGSSAAVAAPLIGLQIMCTAPGAOVGSHWAREAP 240

Db 181 EPGAAAAAVSTYGTGTPFAARGADFOALPVGSSAAVAAPLIGLQIMCTAPGAOVGSHWAREAP 240
QY 241 GAMDCSVENGCEBACNAIPGAPRCQCPAGALADADRSCTASATOSCDNDCEHFCVNP 300
Db 241 GAMDCSVENGCEBACNAIPGAPRCQCPAGALADADRSCTASATOSCDNDCEHFCVNP 300
QY 301 DPGSYSCMCETGYRLADQHRCEVDVDCILBSPCFORCVNTQSGFECCHYPNYDLVNG 360
Db 301 DPGSYSCMCETGYRLADQHRCEVDVDCILBSPCFORCVNTQSGFECCHYPNYDLVNG 360
QY 361 ECVEFVDPCEFRANCEYOCOPINOTSYLCVCAEGFAP1PHEPHRCQMFNCQIACPADCDPN 420
Db 361 ECVEFVDPCEFRANCEYOCOPINOTSYLCVCAEGFAP1PHEPHRCQMFNCQIACPADCDPN 420
QY 421 TQASCEPEGYIIDDGFICTDIDECENGFGCSVCCHNLPGTFECTICGPDALVRIHIGTDC 480
Db 421 TQASCEPEGYIIDDGFICTDIDECENGFGCSVCCHNLPGTFECTICGPDALVRIHIGTDC 480
QY 481 DSGKVD 486
Db 481 DSGKVD 486

RESULT 14

Sequence 1, Application US/07835436C-1
GENERAL INFORMATION:
APPLICANT: NII, ATSUSHI
APPLICANT: MORISHITA, HIDEAKI
APPLICANT: UEMURA, AKIO
APPLICANT: MOCCHIDA, EI
TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIEF & BERRIDGE
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/835,436C
FILING DATE: 26-FEB-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OLIEF, JAMES A.
REGISTRATION NUMBER: 27,075
REFERENCE/DOCKET NUMBER: JAO 27706
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6400
TELEFAX: (703) 836-2787
TELEX: 90-1799 PTO ALEX
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-07-835-436C-1

Query Match 92.2%; Score 2690; DB 3; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.6e-177;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Wed May 8 16:19:25 2002

us-09-509-994-1.1.ram

Page 9

```
QY 1 MGLVYLGLALAGLGFPAAPAPQPGSGQVEHDCFALYPGATFLNASQICDGLRGHLM 60
DB 1 MGLVYLGLALAGLGFPAAPAPQPGSGQVEHDCFALYPGATFLNASQICDGLRGHLM 60
QY 61 TVSSVAADYISLLNGDGVGRRRLMIGLQLPFGCGDPRKRLGFLGFWMTGDNNTSYS 120
DB 61 TVSSVAADYISLLNGDGVGRRRLMIGLQLPFGCGDPRKRLGFLGFWMTGDNNTSYS 120
QY 121 RMARLDLNGAPLPGPLCVASAATVPSEPIWEDQCEVKADGFLCEFHFPATCRPLAV 180
DB 121 RMARLDLNGAPLPGPLCVASAATVPSEPIWEDQCEVKADGFLCEFHFPATCRPLAV 180
QY 181 EPGAAAAVSTIYGTFFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAOGHMAAREAP 240
DB 181 EPGAAAAVSTIYGTFFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAOGHMAAREAP 240
QY 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRCTASATOSCNDLCEHFCVPPNP 300
DB 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRCTASATOSCNDLCEHFCVPPNP 300
QY 301 DPGSYSCMCETGYRLAADQHRCEVDYDCILEPSPCPCRCVNTOGGFECHCYPNYDLVDG 360
DB 301 DPGSYSCMCETGYRLAADQHRCEVDYDCILEPSPCPCRCVNTOGGFECHCYPNYDLVDG 360
QY 361 ECVPEVDPCEFRANCEYOCQPLNOTSYLCVCAEGFAPLPHBPHRCQMFQNTACPADCDPN 420
DB 361 ECVPEVDPCEFRANCEYOCQPLNOTSYLCVCAEGFAPLPHBPHRCQMFQNTACPADCDPN 420
QY 421 TQASCEPBGYIILDDGFICTDIDECENGFGSGVCHNLPGTFECICGPDALVRH 475
DB 421 TQASCEPBGYIILDDGFICTDIDECENGFGSGVCHNLPGTFECICGPDALVRH 475

RESULT 15
US-07-835-436-1
; Sequence 1, Application US/07835436
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFE & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/835,436
; FILING DATE: 19920327
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFE, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-07-835-436-1

Query Match 92.1%; Score 2687; DB 3; Length 475;
Best Local Similarity 99.8%; Pred. No. 2.5e-177;
Matches 474; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLVYLGLALAGLGFPAAPAPQPGSGQVEHDCFALYPGATFLNASQICDGLRGHLM 60
DB 1 MGLVYLGLALAGLGFPAAPAPQPGSGQVEHDCFALYPGATFLNASQICDGLRGHLM 60
QY 61 TVSSVAADYISLLNGDGVGRRRLMIGLQLPFGCGDPRKRLGFLGFWMTGDNNTSYS 120
DB 61 TVSSVAADYISLLNGDGVGRRRLMIGLQLPFGCGDPRKRLGFLGFWMTGDNNTSYS 120
QY 121 RMARLDLNGAPLPGPLCVASAATVPSEPIWEDQCEVKADGFLCEFHFPATCRPLAV 180
DB 121 RMARLDLNGAPLPGPLCVASAATVPSEPIWEDQCEVKADGFLCEFHFPATCRPLAV 180
QY 181 EPGAAAAVSTIYGTFFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAOGHMAAREAP 240
DB 181 EPGAAAAVSTIYGTFFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAOGHMAAREAP 240
QY 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRCTASATOSCNDLCEHFCVPPNP 300
DB 241 GAMDCEVNGGCEHACNAIPGARPCQCPAGALQADGRCTASATOSCNDLCEHFCVPPNP 300
QY 301 DPGSYSCMCETGYRLAADQHRCEVDYDCILEPSPCPCRCVNTOGGFECHCYPNYDLVDG 360
DB 301 DPGSYSCMCETGYRLAADQHRCEVDYDCILEPSPCPCRCVNTOGGFECHCYPNYDLVDG 360
QY 361 ECVPEVDPCEFRANCEYOCQPLNOTSYLCVCAEGFAPLPHBPHRCQMFQNTACPADCDPN 420
DB 361 ECVPEVDPCEFRANCEYOCQPLNOTSYLCVCAEGFAPLPHBPHRCQMFQNTACPADCDPN 420
QY 421 TQASCEPBGYIILDDGFICTDIDECENGFGSGVCHNLPGTFECICGPDALVRH 475
DB 421 TQASCEPBGYIILDDGFICTDIDECENGFGSGVCHNLPGTFECICGPDALVRH 475
```

Search completed: May 8, 2002, 12:43:30
Job time: 341 sec

Gencore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:44:48 ; Search time 48.4 Seconds
(without alignments)
1559.432 Million cell updates/sec

Title: US-09-509-994-2
Perfect score: 2916
Sequence: 1 MGVLVIGALALAGLGFAP.....PSFTPGSTLTPPAVLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_17.*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_fodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-----------|--------------------|
| 1 | 2658 | 91.2 | 468 | 4 Q9UC32 | Q9UC32 homo sapien |
| 2 | 1844 | 63.2 | 577 | 11 Q35370 | P97883 rattus norv |
| 3 | 1400.5 | 48.0 | 461 | 11 P97883 | Q9NPY3 homo sapien |
| 4 | 587 | 20.1 | 652 | 4 Q9NPY3 | Q9HCU0 homo sapien |
| 5 | 584.5 | 20.0 | 757 | 4 Q9HCU0 | Q00274 homo sapien |
| 6 | 584 | 20.0 | 652 | 4 Q00274 | Q08103 mus musculu |
| 7 | 547 | 18.8 | 644 | 11 Q08103 | Q9E161 rattus norv |
| 8 | 534.5 | 18.3 | 643 | 11 Q9ET61 | Q91426 rattus norv |
| 9 | 532.5 | 18.3 | 643 | 11 Q9J126 | Q9TVQ2 caenorhabdi |
| 10 | 431 | 14.8 | 1664 | 5 Q9TVQ2 | Q08281 rattus norv |
| 11 | 414 | 14.2 | 1574 | 11 Q08281 | Q9V589 drosophila |
| 12 | 386 | 13.2 | 1394 | 5 Q9V589 | P87363 gallus gall |
| 13 | 376 | 12.9 | 708 | 13 P87363 | Q9WU19 rattus norv |
| 14 | 375 | 12.9 | 2906 | 11 Q9WU19 | Q08840 mus musculu |
| 15 | 367 | 12.6 | 3857 | 11 Q08840 | Q9WU18 rattus norv |
| 16 | 361 | 12.4 | 2872 | 11 Q9WU18 | Q9CX88 mus musculu |
| 17 | 359 | 12.3 | 528 | 11 Q9CX88 | Q99K58 mus musculu |
| 18 | 358.5 | 12.3 | 1174 | 11 Q99K58 | Q9NSZ1 homo sapien |
| 19 | 357 | 12.2 | 741 | 4 Q9NSZ1 | |

| | | | | | |
|----|-------|------|------|-----------|--------------------|
| 20 | 353.5 | 12.1 | 1242 | 4 Q9NS15 | Q9NS15 homo sapien |
| 21 | 353.5 | 12.1 | 1581 | 4 Q9H7K2 | Q9H7K2 homo sapien |
| 22 | 351.5 | 12.1 | 1312 | 4 Q75412 | Q75412 homo sapien |
| 23 | 351.5 | 12.1 | 1587 | 4 Q00508 | Q00508 homo sapien |
| 24 | 341 | 11.7 | 576 | 4 Q9Y3V7 | Q9Y3V7 homo sapien |
| 25 | 340 | 11.7 | 1062 | 11 Q60789 | Q60789 mus musculu |
| 26 | 333.5 | 11.4 | 961 | 11 Q9E0C6 | Q9E0C6 mus musculu |
| 27 | 332.5 | 11.4 | 1764 | 11 Q35806 | Q35806 rattus norv |
| 28 | 332.5 | 11.4 | 1964 | 11 Q35442 | Q35442 mus musculu |
| 29 | 332 | 11.4 | 999 | 4 Q9NQ36 | Q9NQ36 homo sapien |
| 30 | 331.5 | 11.4 | 1713 | 11 Q08349 | Q08349 mus musculu |
| 31 | 331 | 11.4 | 589 | 5 Q9T2S1 | Q9T2S1 caenorhabdi |
| 32 | 329 | 11.3 | 2189 | 5 Q9B105 | Q9B105 elmeria ten |
| 33 | 328.5 | 11.3 | 956 | 11 Q99K64 | Q99K64 mus musculu |
| 34 | 325 | 11.1 | 1095 | 11 Q60784 | Q60784 mus musculu |
| 35 | 324.5 | 11.1 | 937 | 5 Q9BLJ1 | Q9BLJ1 ciona intes |
| 36 | 323 | 11.1 | 1253 | 11 Q61810 | Q61810 mus musculu |
| 37 | 321 | 11.0 | 2321 | 4 Q9Y6L8 | Q9Y6L8 homo sapien |
| 38 | 321 | 11.0 | 2321 | 4 Q9UM47 | Q9UM47 homo sapien |
| 39 | 319.5 | 11.0 | 1833 | 11 Q00899 | Q00899 mus musculu |
| 40 | 316 | 10.8 | 2281 | 4 Q9NPL3 | Q9NPL3 homo sapien |
| 41 | 314.5 | 10.8 | 3507 | 5 Q23587 | Q23587 caenorhabdi |
| 42 | 314 | 10.8 | 997 | 11 Q9JUS0 | Q9JUS0 mus musculu |
| 43 | 313 | 10.7 | 798 | 5 Q18026 | Q18026 caenorhabdi |
| 44 | 312.5 | 10.7 | 1821 | 4 Q14767 | Q14767 homo sapien |
| 45 | 312 | 10.7 | 517 | 4 Q9NP01 | Q9NP01 homo sapien |

ALIGNMENTS

| | | | |
|---|--------------|------|---------|
| RESULT 1 | PRELIMINARY; | PRT; | 468 AA. |
| Q9UC32 | | | |
| ID Q9UC32: | | | |
| AC Q9UC32: | | | |
| DT 01-MAY-2000 (TREMBLrel. 13, Created) | | | |
| DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update) | | | |
| DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update) | | | |
| DE THROMBOMODULIN. | | | |
| OS Homo sapiens (Human). | | | |
| OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| OX NCBI_TaxID=9606; | | | |
| RN [1] | | | |
| RP SEQUENCE. | | | |
| RX MEDLINE=93293792; PubMed=8390446; | | | |
| RA Yamamoto S., Mizoguchi T., Tamaki T., Ohkuchi M., Kimura S., Aoki N.; | | | |
| RT "Urinary thrombomodulin, its isolation and characterization." | | | |
| RL J. Biochem. 113:433-440(1993). | | | |
| DR HSSP: P07204; 12AO. | | | |
| DR InterPro: IPR00152; Asx_hydroxyl. | | | |
| DR InterPro: IPR000561; EGF-like. | | | |
| DR InterPro: IPR001881; EGF-Ca. | | | |
| DR InterPro: IPR001304; lectin-C. | | | |
| DR InterPro: IPR001491; Thrombomodulin. | | | |
| DR Pfam: PF00008; EGF; 5. | | | |
| DR Pfam: PF00059; lectin.C; 1. | | | |
| DR PRINTS: PR00907; THROMBOMODULN. | | | |
| DR SMART: SM0034; CLECT; 1. | | | |
| DR SMART: SM00181; EGF; 6. | | | |
| DR PROSITE: PS00010; ASX_HYDROXYL; 1. | | | |
| DR PROSITE: PS50041; C-TYPE LECTIN_2; 1. | | | |
| DR PROSITE: PS01186; EGF_2; 2. | | | |
| DR PROSITE: PS01187; EGF_CA; 1. | | | |
| KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat. | | | |
| DR CaScitec: SC01187; EGF-CA; 1. | | | |
| SEQUENCE 468 AA; 49444 MW; 4BFE8E9B8F86A40 CRC64; | | | |

Query Match 91.2%; Score 2658; DB 4; Length 468;
Best Local Similarity 99.8%; Pred. No. 4.4e-204;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 19 APAEPGSGQVEHDCFAFLYGPATFLNASQICDGLRGLMTVRSVADYISLLAND 78

```

Db 1 APAEPQPGSGOCVEHDCALYPGPATFLNASQICDGLGHMTVRSSVAADVISLILNGD 60
QY 79 GGVGRRRLMIGIQLPPGGGDKRLGLPGFQWMTGDNNTSYSRMARLDLNGAPLCGPLCY 138
Db 61 GGVGRRRLMIGIQLPPGGGDKRLGLPGFQWMTGDNNTSYSRMARLDLNGAPLCGPLCY 120
QY 139 AVSAEAATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEGAAAASITYGPFPA 198
Db 121 AVSAEAATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEGAAAASITYGPFPA 180
QY 199 ARGADFQALPVGSSAAVAPLGLQIMCTAPRGAVOGHMAAREAPGAMDCSVENGGEHCNA 258
Db 181 ARGADFQALPVGSSAAVAPLGLQIMCTAPRGAVOGHMAAREAPGAMDCSVENGGEHCNA 240
QY 259 IPGARCCCPAGAAALQADGRCSTASATQSCNDLCEHFCVPMPDQPGSYSCCEFGYRLAA 318
Db 241 IPGARCCCPAGAAALQADGRCSTASATQSCNDLCEHFCVPMPDQPGSYSCCEFGYRLAA 300
QY 319 DOHRCEDVDCTLEBSPCPCRCVNTQSGFECCHYPNVDLVDECEVEPVPCFRANCEYOC 378
Db 301 DOHRCEDVDCTLEBSPCPCRCVNTQSGFECCHYPNVDLVDECEVEPVPCFRANCEYOC 360
QY 379 OPLNOTSYLCVCAEGFAPIPEHPRHRCMFCNOTACPADCDNTQASGCEPFGYTLDDGFI 438
Db 361 OPLNOTSYLCVCAEGFAPIPEHPRHRCMFCNOTACPADCDNTQASGCEPFGYTLDDGFI 420
QY 439 CTDIDECENGSGFSCVCHNLPGTFECICGPPSALAHNIGTDCDSKXVD 486
Db 421 CTDIDECENGSGFSCVCHNLPGTFECICGPPSALAHNIGTDCDSKXVD 468

RESULT 2
035370 PRELIMINARY; PRT; 577 AA.
AC 035370:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE THROMBOMODULIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
RA Wang J., Yao A., Wang J.Y., Hardin J.W., Hauer-Jensen M.;
RN Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 1-136 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Yao A., Wang J., Hardin J.W., Hauer-Jensen M.;
RN Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF022743; AAB80760.1; -
DR EMBL; AF022742; AAB80923.1; -
DR HSSP; P07204; IFGD.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001304; Ictln_C.
DR InterPro; IPR001491; Thrombomoduln.
DR Pfam; PF00008; EGF_5.
DR PRINTS; PR00907; THROMBOMODULN.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA_2.
DR SMART; SM00001; EGF_Like; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA_2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.

```

```

SQ SEQUENCE 577 AA; 61844 MW; 0BE764C8BE18555F CRC64;
Query Match 63.2%; Score 1844; DB 11; Length 577;
Best Local Similarity 64.0%; Pred. No. 4, 1e-139;
Matches 333; Conservative 49; Mismatches 132; Indels 6; Gaps 5;
QY 1 MGVVLVGLALATLGFPAPEPQPGSGOCVEHDCALYPGPATFLNASQICDGLGHLM 60
Db 1 MGVVLVGLALATLGFPAPEPQPGSGOCVEHDCALYPGPATFLNASQICDGLGHLM 60
QY 61 TVSSVAADVISLILNGDGGVGRRLMIGIQLPPGGGDKRLGLPGFQWMTGDNNTSY 120
Db 61 TVSSVAADVISLILNGDGGVGRRLMIGIQLPPGGGDKRLGLPGFQWMTGDNNTSY 118
QY 121 RMARLDLNGAPLCPGVAVSAEAATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAV 180
Db 119 RMARLDLNGAPLCPGVAVSAEAATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAV 178
QY 181 EP-GAAAASITYGPFPAARGADFGALPVGSSAAVAPLGLQIMCTAPRGAVOGHMA 239
Db 179 NTRDPEGAHISSTYNTPLVGSADFOTLPIGSAATVAPGELVCRALPGTSEGHWTREV 238
QY 240 PGAMDCSVENGGEHCNAIPGARCCCPAGAAALQADGRCSTASATQSCNDLCEHFCVPN 299
Db 239 PGAMDCSVENGGEHCNAIPGARCCCPAGAAALQADGRCSTASATQSCNDLCEHFCVPN 298
QY 300 PDQPGSYSCCEFGYRLAADQHRCEDVDCTLEBSPCPCRCVNTQSGFECCHYPNVDLV 359
Db 299 SDVPGSYSCCEFGYRLAADQHRCEDVDCTLEBSPCPCRCVNTQSGFECCHYPNVDLV 358
QY 360 GECVEPVPCFRANCEYOCPLNOTSYLCVCAEGFAPIPEHPRHRCMFCNOTACPADCDP 419
Db 359 GECVEPVPCFRANCEYOCPLNOTSYLCVCAEGFAPIPEHPRHRCMFCNOTACPADCDP 418
QY 420 NTQASCEPFGYTLDDGFICTDIDECENGSGFSCVCHNLPGTFECICGPPSALAHNIG 479
Db 419 NTQASCEPFGYTLDDGFICTDIDECENGSGFSCVCHNLPGTFECICGPPSALAHNIG 478
QY 480 CDSGKY--DGDGSGSGEPPEPS-PTPGSTLPPAVGLVHSG 516
Db 479 CDPVLEDESDGSGEPHSSNTPTVSSITVPPARMHSG 518

RESULT 3
P97883 PRELIMINARY; PRT; 461 AA.
AC P97883:
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE THROMBOMODULIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN CAPILLARY;
RA Wang L., Tran N.D., Schreiber S.S., Zlokovic B.V.;
RN Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90121; AAB49723.1; -
DR HSSP; P07204; IFGD.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001491; Thrombomoduln.
DR Pfam; PF00008; EGF_5.
DR PRINTS; PR00907; THROMBOMODULN.
DR SMART; SM00179; EGF_CA_2.
DR SMART; SM00001; EGF_Like; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS01186; EGF_2; 3.

```


RL J. Biol. Chem. 276:7408-7414(2001).
 DR EMBL; AF279142; AAG00867.1; -
 DR EMBL; AJ295846; CAC34381.1; -
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF-Ca.
 DR InterPro; IPR001304; lectin_c.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00059; lectin_c; 1.
 DR SMART; SM00034; CLECT; 1.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00179; EGF_Ca; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
 KW Signal.
 FT SIGNAL. 1 17
 FT CHAIN 18 757 TUMOR ENDOTHELIAL MARKER 1.
 SQ SEQUENCE 757 AA; 80859 MW; C96363EAI1FD8FFA0 CRC64;

Query Match 20.0%; Score 584.5; DB 4; Length 757;
 Best Local Similarity 33.6%; Pred. No. 1,3e-38;
 Matches 177; Conservative 44; Mismatches 191; Indels 115; Gaps 24;

QY 1 MGVLVGLALALAGLGFAPAPAPPGSGQVCHDCEPALTGPATFLNASQICDGLRGHL 59
 DB 2 LKRLLLAMAAAGPTLGDPPAAAPR--AACGPGSCYALFPRRTTLEAMRACRELIGDL 58
 QY 60 MYRVSSVADVISLLNDGSGVRRRLMIGLPPGCGDPKRLGRLPGFWMTGDNNTSY 119
 DB 59 APTPTTEEAQRNDISLVG--AGPASKRLIMIGLQARQCOLR--PLKGFMTGDDOTAR 114
 QY 120 SRMARLDLNGAPLGLPLCYAASAETVSEPTWEEQCEVKADGELCEHPFATCPRLA 179
 DB 115 TWWAQ-PASGGPCPAPORCYALEAS--GEHRLGSGCLTAVDYLQCGFEFGACAPALQ 168
 QY 180 VEPGAAAAVSTIYGPFPARGADPQALRYGSSAAV--APGLQIMCTAPGAVGHHMA 236
 DB 169 DEAGQAGPAPV--YTPPHLVSTEFEMLPFGSYAAVQCAGRGASLICYAPREGVGW-MS 224
 QY 237 REAP--GAMDCSVENGGEHAC-NAIPGAPRCQAGALADGSSCTASATQS-CNDL 291
 DB 225 RAGPLCLGT-GSPDNGGGEHECEVEVDHVSCTEGEGRILADGSSCEPCQAACDEQ 283
 QY 292 CEHFVCPNDPQGSYSQCMETGYRLAADQ-HRCEDVDCCILEPSPQRCVNTQGGFECH 350
 DB 284 CE-----PGGPGQYSCHRLGFRPAEDDPHRCVDTDEQI-AGVQOQCVNTVGGFECH 336
 QY 351 CYPNVDLVGCEVEPVPDPCFRANCEYQCPLNQSTYLCAEGFADIPHEPHRCQMFNQ 410
 DB 337 CSEGHLE-----ADGIS----- 349
 QY 411 TACPADCDP-----NQASCEPEGYILLDDGFICTDDE--GENGFGSGVCHNLPGT-- 461
 DB 350 -----CSPAGMAGQASQDLDG-ELDDGDEDEDEEMKAFNGMT-----EMPGILM 397
 QY 462 FEGICGSPSALARHIGTDCDSKVDGSGSEPPSPPGSTLMP 508
 DB 398 MEPTQPPDFALAYRPSFPD-----REFQIPIPEPTWPP 432
 RESULT 6
 ID 000274 PRELIMINARY; PRT; 652 AA.
 AC 000274;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE C10R.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 QX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97199258; Pubmed=9047234;
 RA Neomunceno R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;
 RT "cDNA cloning and primary structure analysis of C1q(RP), the human
 C1q/MBL/SPA receptor that mediates enhanced phagocytosis in vitro."
 RL Immunity 6:119-129(1997).
 DR EMBL; U94333; AAB53110.1; -
 DR HSSP; P35555; IEMN.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001304; lectin_c.
 DR Pfam; PF00008; EGF; 5.
 DR Pfam; PF00059; lectin_c; 1.
 DR SMART; SM00034; CLECT; 1.
 DR SMART; SM00179; EGF_Ca; 3.
 DR SMART; SM00001; EGF_Like; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 3.
 DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
 DR PROSITE; PS01186; EGF; 2; 3.
 DR PROSITE; PS01187; EGF_Ca; 3.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 SQ SEQUENCE 652 AA; 68576 MW; B7EAB5FE5714A775 CRC64;

Query Match 20.0%; Score 584; DB 4; Length 652;
 Best Local Similarity 31.8%; Pred. No. 1,2e-38;
 Matches 181; Conservative 56; Mismatches 209; Indels 124; Gaps 32;

QY 2 LGVLVGLALALA--GLGFAPAPAPPGSGQVCHDCEPALTGPATFLNASQICDGLRGHL 59
 DB 5 MGILLILLILLIPRGAGTGADTE-----AYVCGTACTYTHSGLSAAEAONHCNONGNL 60
 QY 60 MYRVSSVAD--VISLLNDGSGVRR--RLMIGLQPPG-CGDPKRLGRLPGFWMT 112
 DB 61 ATVKSKEEAQHYORVLAQLRLREALTARMSKFTWIGLQREKGCIDPSL--PLKGSWVG 118
 QY 113 GDNNTYSRRARLDLNGAPLGLPLCYA--VSAEAATVPS-PIWEEQOC-----EVKAD 163
 DB 119 GGEEDTYSNMHKLRLNSC--ISKRCVSLDLSQPLLPNLPKMSGSPGSGPSPSINIE 176
 QY 164 GFLCEHPFATCPRLAV-EPGAAAAVSTIYGPFPARGADPQALRYGSSAAVAPLGLQL 222
 DB 177 GFYCKSFQKMCNPLALGPG-----OYITTFQTTSSLEAVFPASANA----- 224
 QY 223 MC-TAPGAVGHW--AREAPGAND-----CSYENGCEHAC-NAIPGAP 263
 DB 225 -CGEEDKDETQSHYFLCKEKAPDVPFDMGSSGPLCYPKKGCNNGNCGHODCEGGDGSF 283
 QY 264 KCQCPGAGALADGRCSTASATQSCNDLCE--HFCVPNDPQGSYSQCMETGYRLAADQ 321
 DB 284 LCGCRPGFRLDLDLVTC-ASRNPCSSSPCRGATCVLGP-HGKNYTCRCPOGYLDSQL 341
 QY 322 RCEDVDCCILEPSPQRCVNTQGGFECHCYRYDLYDCEVPPVPCFRANCEYQCP 381
 DB 342 DCYVDVDEC--QSPCAQCECVNTPGFRCECWGY-----EPGSP-----GEGACQV 386
 QY 382 NQTSYLCVCAEGFAPRPHRPHRCQMFQNTACPADCDPNTQAS--CECEGYIL--DDGF 437
 DB 387 DE-----CALGRSP-----CAQGC-TTIDGSPFCSEGYVLAGEDGT 423
 QY 438 ICTDIDEC--ENGFGSGVCHNLPGTEFCICGPDALARHIGTDCDSKV-----D 486
 DB 424 QCDVDCEVGPGLCDLSCFTNQGSGFHCGLPGWVLAN-GVSCMTMGVSLGPPSGPPD 482
 QY 487 GGDSSGE-----PPSPPTPGSTLTPPA 509
 DB 483 EEDKGEKESIVPRAATASPTRGPGETPKA 512

RESULT 7
 089103


```

ID 089103      PRELIMINARY:      PRT:      644 AA.
AC 089103;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE C1Q/MBL/SPA RECEPTOR C1QRP.
GN LY68 OR C1QRP OR A44.
OS Mus musculus (Mouse).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SEV;
RA Tenner A.J., Kim T.S.;
RT "Identification of the mouse genomic DNA for C1QRP,"
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Pettenko O., Lemischka I.R.;
RT "Molecular characterization of A44, an early marker of hematopoietic
development,"
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SEV;
RX MEDLINE=99359842; PubMed=10430665;
RT Norsworthy P.J., Taylor P.R., Walport M.J., Bolto M.;
RT "Cloning of the mouse homolog of the 126-KDa human C1q/MBL/Sp-A
receptor, C1QRP,"
RL Mamm. Genome 10:789-793(1999).
DR EMBL: AF074856; AAC63274.1;
DR EMBL: AF081789; AAC62649.1;
DR EMBL: AF099939; AAD47906.1;
DR EMBL: AF099938; AAD47906.1; JOINED.
DR HSSP: P07204; 1EGT.
DR MDG: MG110666; Ly68.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_5.
DR SMART: SM00059; lectin_c; 1.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00179; EGF_Ca; 3.
DR SMART: SM00001; EGF_Like; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 3.
DR PROSITE: PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_Ca; 3.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
KW Receptor; Repeat.
SQ SEQUENCE 644 AA; 69354 MW; E84351648BF86354 CRC64;

```

```

Query Match      18.8%; Score 547; DB 11; Length 644;
Best Local Similarity 30.2%; Pred. No. 1,1e-35;
Matches 168; Conservative 55; Mismatches 210; Indels 124; Gaps 27;

```

```

DB 184 KMCRLPLGLGPG-----RVYTTPTPOATSSLEAVPFAVANVA-CGDGAKSETHYFL 236
OY 224 CP-APPGAVOGHWAREAD---GAMDCSVENGGEHAC-NAIQAPRCQCPAGALQADG 277
DB 237 CNEKTRGIF--HMGSGLCVSPKFGCSFNNGGCGQDCFGGDSFGRCGRPGFRLDDL 294
OY 278 RCTASATQSCNDLCEHFCVNPDPG-----SYGKETEGRCLAADHREDVD 327
DB 295 VYCA-----RNCSSNPCTGGCMCHSVPLSENRYCRCPGQYQDSSQVHCVDI 344
OY 328 DCILPPSPCPQCVNTQGGFCHCYPNVDLVG---ECVPPVPCFRANCYOCPLNOT 384
DB 345 EC--QDSFCADQCVNTGLSFHCECWVGQ--PSGKEEACEDVDECAANSP----- 392
OY 385 STLYCAGSGFADIPHEPRCOMFCNQTACPADCPNTQAS--CECPGYTL--DDGFICT 440
DB 393 -----CAQGCI-----NTDGSFYCSCKEGYIVSGDSSTOGE 423
OY 441 DIDECEN--GGFCGCHNLPGTFECIGPDSALARIHGTCDGSG-----KYD 486
DB 424 DIDECSDARGNRCDSLCNTDGSFRCGCPQEMELAPN-GVFCGRGYFSLPARPQKED 482
OY 487 GSDSGSGEPPSPSPG 503
DB 483 NDRKESTWPTPEMPSS 499

```

```

RESULT 8
O9ET61
ID 09ET61      PRELIMINARY:      PRT:      643 AA.
AC 09ET61;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE C1Q/MBL/SPA RECEPTOR C1QRP PRECURSOR.
GN C1QRP.
OS Rattus norvegicus (rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PGV;
RA Lovik G., Vaage J.T., Disen E., Szpirer C., Ryan J.C., Rolstad B.;
RT "Characterization and molecular cloning of rat C1qR, a receptor on
macrophages, natural killer cells and dendritic cells,"
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF136537; AAG01572.1;
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_5.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF_5.
DR SMART: SM00179; EGF_Ca; 5.
DR SMART: SM00001; EGF_Like; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN.1.
DR PROSITE: PS00041; C_TYPE_LECTIN_2; 1.
KW Signal; Receptor.
FT SIGNAL 1
SQ SEQUENCE 643 AA; 68781 MW; 9A64C933AD943DB6 CRC64;

```

```

Query Match      18.3%; Score 534.5; DB 11; Length 643;
Best Local Similarity 29.9%; Pred. No. 1,1e-34;
Matches 167; Conservative 56; Mismatches 206; Indels 127; Gaps 28;

```

```

OY 4 VIVLGAIALAGLGFPAPEPOGSGCYEHDCFALYPGPATFLNASOICDGLRGLMTVR 63
DB 8 LILGLIGLQMGAAADSE---AVVCEGTACTYTAHMGKLSAFAOHRCNENGNGLATYK 63
OY 64 SSVA-----DIVISLLN-----GDGVRGRRLMIGLQLPFGCGDKRRLGRLPGFWTVGDN 115

```

DB 64 SEERHVOEALADLLKTKKASEETKIG--KFWIGLOREKCKTYHDL-PMKGF5WVGGE 120
 OY 116 NTYSRMAWRLDNG--APLCGPLCVAASAATVPSE-PIWEEQOCEV-----KADGFL 166
 DB 121 DTTYSNMWTKASKSCISKRCVSLIDLSLKP--PSHLPRMHSRPGCTDAPAGNSTEGFL 178
 OY 167 CEFHPATCRPLAV-EPGAAAASVITYGTPFAAGADFOALPYGSSAAVAPLGIQ---- 221
 DB 179 CKFNKMGCSPLALGPG-----QLTYTTFQATTSLSKAVFASVANYV-CGDEAESK 231
 OY 222 ---LMCTAPGAVOGHWAAREP---GAMDCSVENGSCENAC-NALPGAAPROCPAGAL 273
 DB 232 TNYLCKRTTAGV-FHWSSGSLCVSPKFGCSFNNGCCODCEGGDGSFRGCGRPRFL 290
 OY 274 QADGSCSTASATQSCNDLCEHFCVNPDPG-----SYSCMCEGTGRLAADQHR 323
 DB 291 LDDLVTGAS-----RNPCCSNPCTGGMCHSVPLSENCTHCPRGYQLDSQVHC 340
 OY 324 EDVDDCILEPSPORCVNTGSGFCHCYPNYDLVDG--ECVPEVDPGFRANCEYQOCP 381
 DB 341 VDIDEC--EDSPCDOECINTPGFHCRCWGYOSSGSKERACEDVDEC----- 386
 OY 362 NOTSYLCVCAEGFAPIPHEPRHRCMFCNQACPADCDPNTQAS--CECEGYIL--DDGF 437
 DB 387 --TAAYSPCAOG-----CT-----NTDGSFYSCKEGYIMSGEDST 420
 OY 438 ICTDIDECENGSGCVCHNLPGTFECICGPDALAHIGTDCDSG-----KV 485
 DB 421 QCEDIDEC-IGNPCDTLICINTDGSFRGCGPAGFELAPN-GVSCTRGSMFSELPARPOKE 478
 OY 486 DSGDSGSGEPPSPPTGS 503
 DB 479 DKDGKESTVPLTEMPGS 496

RESULT 9
 O9J126 PRELIMINARY; PRT: 643 AA.
 AC O9J126 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE CLOP.
 GN CLOP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN-MISTER: TISSUE-LUNG;
 RT Dean Y.D., McCreath E.P., Akatsu H., Morgan B.P., Gasque P.;
 RT "Cloning of the Rat CjgR cDNA";
 RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF160978; AF80402.1;
 DR InterPro: IPR000152; ASX_hydroxyl.
 DR InterPro: IPR000561; EGF_Ca.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001304; lectin_C.
 DR Pfam: PF00008; EGF_5.
 DR SMART: SM00034; CIECT; 1.
 DR SMART: SM00179; EGF_Ca; 3.
 DR SMART: SM00001; EGF_Like; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 3.
 DR PROSITE: PS00041; C-TYPE_LECTIN_2; 1.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_Ca; 3.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 SQ SEQUENCE 643 AA; 68780 MW; FB23D742E02A4931 CRC64;

Query Match 18.3%; Score 532.5; DB 11; Length 643;

Best Local Similarity 29.9%; Pred. No. 1,5e-34;
 Matches 167; Conservative 55; Mismatches 209; Indels 127; Gaps 28;
 OY 4 VIVIGLALAGGCPAPAPQPGSGQCEHDCPFLYGPATFLINASQIDGGLGHMTVR 63
 DB 8 LILIGLIGLWMAAARDE----AVCEGTACTAHMGKLSAABDRHRCNENGSLATYK 63
 OY 64 SSVAA---DVISLLN---GDGSGRRRLWIGLQIPGCGDPKRLGFLRGFWYTGDN 115
 DB 64 SEERHVOEALADLLKTKKASEETKIG--KFWIGLOREKCKTYHDL-PMKGF5WVGGE 120
 OY 116 NTYSRMAWRLDNG--APLCGPLCVAASAATVPSE-PIWEEQOCEV-----KADGFL 166
 DB 121 DTTYSNMWTKASKSCISKRCVSLIDLSLKP--PSHLPRMHSRPGCTDAPAGNSTEGFL 178
 OY 167 CEFHPATCRPLAV-EPGAAAASVITYGTPFAAGADFOALPYGSSAAVAPLGIQ---- 221
 DB 179 CKFNKMGCSPLALGPG-----QLTYTTFQATTSLSKAVFASVANYV-CGDEAESK 231
 OY 222 ---LMCTAPGAVOGHWAAREP---GAMDCSVENGSCENAC-NALPGAAPROCPAGAL 273
 DB 232 TNYLCKRTTAGV-FHWSSGSLCVSPKFGCSFNNGCCODCEGGDGSFRGCGRPRFL 290
 OY 274 QADGSCSTASATQSCNDLCEHFCVNPDPG-----SYSCMCEGTGRLAADQHR 323
 DB 291 LDDLVTGAS-----RNPCCSNPCTGGMCHSVPLSENCTHCPRGYQLDSQVHC 340
 OY 324 EDVDDCILEPSPORCVNTGSGFCHCYPNYDLVDG--ECVPEVDPGFRANCEYQOCP 381
 DB 341 VDIDEC--EDSPCDOECINTPGFHCRCWGYOSSGSKERACEDVDEC----- 386
 OY 362 NOTSYLCVCAEGFAPIPHEPRHRCMFCNQACPADCDPNTQAS--CECEGYIL--DDGF 437
 DB 387 --TAAYSPCAOG-----CT-----NTDGSFYSCKEGYIMSGEDST 420
 OY 438 ICTDIDECENGSGCVCHNLPGTFECICGPDALAHIGTDCDSG-----KV 485
 DB 421 QCEDIDEC-IGNPCDTLICINTDGSFRGCGPAGFELAPN-GVSCTRGSMFSELPARPOKE 478
 OY 486 DSGDSGSGEPPSPPTGS 503
 DB 479 DKDGKESTVPLTEMPGS 496

RESULT 10
 O9TYO2 PRELIMINARY; PRT: 1664 AA.
 AC O9TYO2 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE Y64G10A.7 PROTEIN.
 GN Y64G10A.7
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN 11
 RP SEQUENCE FROM N.A.
 RA Mortimore B.J.;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN 121
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology";
 RT Science 282:2012-2018(1998).
 RN 131
 RP SEQUENCE FROM N.A.
 RA Ainscough R.;
 RA Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL117206; CAB60454.1; -;

| | | | |
|--------|---|--|---------------|
| OY | 119 | YSRRARLDINGAPLCPGLCAVAAAE-----ATVPSEPIWEQOCF--VRADGFLCFE | 169 |
| Db | 56 | YLRFAFRSRRGCSKCCLLIRQAANCADLCNGGTCTVPSEHNDNEQYCECPVGFYGAQCOY | 115 |
| OY | 170 | HEPATCRPLAVEPGAAAAAVSTTYGTFEFAARGADFOALPVGSSAAVAPLGLQLMCTAPRG | 229 |
| Db | 116 | D-ANECC--MANNGCCHEECVN-TLGTYY-----CRCWPG | 145 |
| OY | 230 | AVQGHMARREAPGAMDCSVEENGCEHNCNAIPGAPROCPAGALADGRSCTA----- | 282 |
| Db | 146 | FELSGDGTCSDDIDECVANSNGGSDRCVNSPGFRCQCPBDLYLHADGRTCGSGHFENL | 205 |
| OY | 283 | ---SATOSC---NDLCHEFCVPNPDGYSKCMCEGYRLAADQHRCEYDDCILEPSPC | 336 |
| Db | 206 | ILIKRVYSCSTNDNGCCHEEC--ENDSNGEFYRCRCRVGFKLSEKRCSCQPDPCFDNRKGC | 264 |
| OY | 337 | PQRCVNPQGGEGCHCYANNYL--VDGECVEYVDQCFRAN--CEYOCQPLNQTSTLYCAAEF | 394 |
| Db | 265 | QHCTNHNHGRACQCCYGFHLSYDRNSCVDDIDECANNGGCEHFCENVKGT-YKCKREGY | 323 |
| OY | 395 | APIDHEHRKCMF-----CMQTAQCPADC--DPNTQAQSCCECPREGYII-DGFICTDIDEC- | 445 |
| Db | 324 | -QLGRDGTCTCEMLGAGCGOVGGNGGCHDQDQPDGSHVCKCRNGYITLADQKLCIHDINECH | 382 |
| OY | 446 | ENGFGCSGYCNLNGTGTCTCIGPDSALARIIGT-----DDDS-----GKVDGSDSG | 491 |
| Db | 383 | ENNDDCSDICVNLGVSVECOCKPGEFLMKRKCTCEDISECSSNNGGCEQJCSNQSGGYMC | 442 |
| OY | 492 | SGEP | 495 |
| Db | 443 | SCEP | 446 |
| RESULT | 11 | | |
| O88281 | | | |
| ID | O88281 | PRELIMINARY; | PRT; 1574 AA. |
| AC | O88281 | | |
| DT | 01-NOV-1998 (Tremblrel, 08, Created) | | |
| DT | 01-NOV-1998 (Tremblrel, 08, Last sequence update) | | |
| DT | 01-JUN-2001 (Tremblrel, 17, Last annotation update) | | |
| DE | MEGF6. | | |
| GN | MEGF6. | | |
| OS | Rattus norvegicus (Rat). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. | | |
| OX | NCBI_TaxID=10116; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN; | | |

| | | | | |
|-----------------------|------------------|------------------|------------|--------------|
| Query Match | 14.28; | Score 414; | DB 11; | Length 1574; |
| Best Local Similarity | 34.18; | Pred. No. 1e-24; | | |
| Matches 104; | Conservative 28; | Mismatches 101; | Indels 72; | Gaps 17; |

RESULT 12
9VS89
ID 09VS89
PRELIMINARY;
PRT; 1394 AA.

DP 01-MAY-2000 (Tremblurel. 13, Created)
 DT 01-MAY-2000 (Tremblurel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblurel. 17, Last annotation update)
 DE CG7526 PROTEIN (FRAGMENT).
 GN CG7526.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celitker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Tandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

```

Query Match Length      13.2%  Score 386; DB 5; Length 1394;
Best Local Similarity 32.1%  Pred. No. 1,66-22;
Matches 99; Conservative 33; Mismatches 110; Indels 66; Gaps 18;

OY 224 CTAPPGAVG-----HAAEPAGM-----DSYENGCEHACNARI6 261
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 563 CLCPGATGLDNHIVTSLNSSTFDSTSTSEPTSHATCLDIDECISLNGMSHRCONEP6 622
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 262 APRCQCPAGAAIQADRSCTASATOSC--NDLCEHFCVPNPDPGASGYSCETGYRLAA 318
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 623 GFQCCADPLGALSEDRKTC--QDIBECILDSNGQCSQLCL--NPGGFACACEGTGFLTP 677
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 319 DQHCCEPVDCLIEPRPCQRCBQCNQGGEGHCYCYNVLYDGE--CYEPVDPG---FRAN 373
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 678 DFGECADIDRCSDYNGCSDICINLGTACACCEGYTGLADKRLSCLD--VDECAGLISLG 736
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 374 CEYQCGPDLNQT--SYLCVCAEGFAPRIIPHEPHRCQMFQCNQTACPA--DCDPNTQASC----- 425
   | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

| RESULT | 13 | | |
|--------|--|--------------|--------------|
| ID | P87363 | PRELIMINARY: | PRT: 708 AA. |
| AC | P87363; | | |
| DT | 01-MAY-1997 (TREMBLrel. 03, Created) | | |
| DT | 01-MAY-1997 (TREMBLrel. 03, Last sequence update) | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last annotation update) | | |
| DE | FIBRILIN-1 (FRAGMENT). | | |
| GN | FN1. | | |
| OS | Gallus gallus (Chicken). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; | | |
| OC | Gallus. | | |
| OX | NCBI_TaxID=9031; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RA | Zhou G., Price C., Godfrey M.; | | |
| RL | Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases. | | |
| DR | EMBL; U088872; AAA8531.1; .. | | |
| DR | HSP; P07204; IFGD. | | |
| DR | InterPro: IPR000152; Asx_hydroxyl. | | |
| DR | InterPro: IPR000561; EGF-like. | | |
| DR | InterPro: IPR001881; EGF_Ca. | | |
| DR | InterPro: IPR002212; TB. | | |
| DR | InterPro: IPR001491; Thrombomoduln. | | |
| DR | Pfam; PR00008; EGF_14. | | |
| DR | Pfam; PF00683; TB_2. | | |
| DR | PRINTS; PRO0907; THROMMODULN. | | |
| DR | SMART; SM00179; EGF_CA; 14. | | |
| DR | PROSITE; PS00010; ASX_HYDROXYL; 13. | | |
| DR | PROSITE; PS01186; EGF_2; 10. | | |
| DR | PROSITE; PS01187; EGF_CA; 13. | | |
| KW | Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat. | | |
| Et | NON_TER 708 | | |
| SEQ | SEQUENCE 708 AA; 76163 MW; C247271CIDF73361 CRC64; | | |

[illegible]

Db 503 RSGGCTDVNCEADPTICISTGCVTAGSYTCECPD-----FELNPRVGCY 550

RESULT 14
Q9WUH9 PRELIMINARY; PRT: 2906 AA.
AC Q9WUH9:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE FIBRILIN-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Ota K., Tian Y., Kumar A., Wada J., Kashihara N.,
RA Wallner E.I., Kanwar Y.S.;
RT "Cloning of rat fibrillin-2 cDNA and its role in branching
morphogenesis of rat embryonic lung."
RL Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF135060; AAD34439.1; .
DR HSSP; P35555; 1EMN.
DR InterPro: IPR002086; Aldehyde_dehydrt.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR002212; TB.
DR Pfam: PF00008; EGF; 46.
DR Pfam: PF00683; TB; 9.
DR PRINTS; PR00010; EGFBL00D.
DR SMART; SM00179; EGF_CA; 42.
DR SMART; SM00001; EGF_Like; 4.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 36.
DR PROSITE; PS01187; EGF_CA; 43.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 2306 AA; 313372 MW; 9EB64E727044EF58 CRC64;

Query Match 12.9%; Score 375; DB 11; Length 2906;
Best Local Similarity 32.9%; Pred. No. 2,4e-21;
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

QY 244 DCSVNGGCEHACNAIPGAPRCQCPAGALADGRCSTASATQSCNDLCEHCVNPND- 301
DB 1237 EGMINGGCDPQCTNSESGSEGYALMPDGRSCA-----DIDE--CENNPDIIC 1286
QY 302 -----QPGSSCMCEGRIADQHRCEVDYDCILPSPCP-QRCVNTGGRECHYR 353
DB 1287 DGGCTNIPGEYRCLCYGPFASMDKTCIDVNECDLNPNCMFGECEHTKGSFCHQQL 1346
QY 354 ATDLVGE--CVEPDPD--FRANCEYQCPIN-OTSYLCVCAEFA-----PIR 398
DB 1347 GYSVKKATGCTD-VDEBEIGHNCDMAASCLINVGSKSCRSRWANGIKICIDLBACA 1405
QY 399 HEPRHCQMFQNTACPADCPNTQAS--CECEGYILDDGFCTDIDE-----CENG- 448
DB 1406 NCTHCCSI-----NAOC-VNTPGSYRCACSEGF-TGDGFTCSVDDECAENINLCENGQ 1456
QY 449 -----GF-----GVCNMLPGFEFCICGPDASALA 473
DB 1457 CLNVGATYCECEMGFTTASDSRSCQDIDECFQNTCVFGTCNNLPGHFCICDDGGLD 1516
QY 474 RHIG--TDGD 481
DB 1517 RTGGHCTDID 1526

RESULT 15
O88840 PRELIMINARY; PRT: 3857 AA.
ID O88840;
AC O88840;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MUTANT FIBRILIN-1.
GN FBN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B10.D2;
RX MEDLINE-98069008; PubMed-9405934;
RA Bona C.A., Mural C., Casares S., Kasturi K., Nishimura H., Honjo T.,
RA Matsuda F.;
RT "Structure of the mutant fibrillin-1 gene in the tight skin (TSK)
mouse."
RL DNA Res. 4:267-271(1997).
DR EMBL; AF007248; AAC62317.1; .
DR HSSP; P35555; 1ABV.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR002557; Chitin-binding.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR002212; TB.
DR InterPro: IPR000822; ZnF-C2H2.
DR Pfam: PF00008; EGF; 64.
DR Pfam: PF00683; TB; 12.
DR SMART; SM00494; ChIPD2; 2.
DR SMART; SM00179; EGF_CA; 60.
DR SMART; SM00001; EGF_Like; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 61.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 50.
DR PROSITE; PS01187; EGF_CA; 61.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 3857 AA; 418303 MW; 5BC061B8C527E04C CRC64;

Query Match 12.6%; Score 367; DB 11; Length 3857;
Best Local Similarity 30.3%; Pred. No. 1,4e-20;
Matches 118; Conservative 28; Mismatches 132; Indels 112; Gaps 21;

QY 212 SAAVAPLGT-----QIMCTAPPGAVOGHWAREPAGMDCSVNGGCEHACNAIFGAPRC 265
DB 2148 SANLCFPHGRVNLIGKYQACNPGYHPTDRFLFCVIDIDECISINNGCEFFCTNSDGYEC 2207
QY 266 QCPAGALADGRCSTASATQSCNDLCEHCVNPND-----DQPGYSQMCETGR 316
DB 2208 SCQPGFALMPDGRSCF-----DIDE--CEDNPICDGGCTNIPGEYRCLCYDGFA 2257
QY 317 AAQNHCEEDVDCILPSPCP-PQRCVNTGGRECHYRPNYDLVDEG--CVEPDPD--FR 371
DB 2258 SEDMKTCVYDNECDLNPNCISGTCENTKGSFCHCDMKYSKKKGTGCTD- INCEFCIGA 2316
QY 372 ANCEYQCPINOT-STLYCVCAEFA-----PIRPHRCQMFQNTACPADCPN 420
DB 2317 HNDDBRAVNTAGTSFKSCSPGMDIGIKYCTDIDECSSGTFHMCSDH-----ADC-KN 2368
QY 421 TQAS--CECEGYILDDGFCTDIDEC-EN-----GGF----- 450
DB 2369 TMGSYRCLCKDGI-TGDGFTCTDIDECSENILNCGGQCLNAPGGYRCRCMDGFPSPADG 2427
QY 451 -----CS-----GVCNMLPGFEFCICGPDASALAHIG-----TDGDSGK 484
DB 2428 KACEDIDECSLPNICVFGTCNNLPGHFCRCCEIGYELDRSGGCTDVNCEADPTICISGN 2487

QY 485 -VDGDSGSGEPPPTPGSTLTPPAVGLV 513
! : : ! ! ! ! !
Db 2488 CVNTPGSYTCDCPPD---FELNPTRVGCV 2513

Search completed: May 8, 2002, 12:44:52
Job time: 353 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:40:00 ; Search time 29.56 Seconds
(without alignments)
1329.705 Million cell updates/sec

Title: US-09-509-994-2

Perfect score: 2916
Sequence: 1 MGVYVIGALALAGLGFAP.....PSPYRGSITLPPAVGLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_68:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|---------------------|
| 1 | 2912 | 99.9 | 575 | 1 | THHUB |
| 2 | 1903 | 65.3 | 577 | 2 | thrombomodulin pre |
| 3 | 1095.5 | 37.6 | 356 | 2 | thrombomodulin - b |
| 4 | 414 | 14.2 | 1574 | 2 | MEGF6 protein - ra |
| 5 | 414 | 14.2 | 1620 | 2 | hypothetical prote |
| 6 | 374 | 12.8 | 2907 | 2 | fibriillin-2 precu |
| 7 | 368 | 12.6 | 2871 | 2 | fibriillin I - bovl |
| 8 | 367.5 | 12.6 | 1184 | 2 | fibriillin-2 precu |
| 9 | 366 | 12.6 | 2918 | 2 | A55184 |
| 10 | 364 | 12.5 | 3002 | 2 | A54105 |
| 11 | 362 | 12.4 | 2871 | 2 | A47221 |
| 12 | 357 | 12.2 | 741 | 2 | A55624 |
| 13 | 357 | 12.2 | 1221 | 2 | A46488 |
| 14 | 352.5 | 11.4 | 1964 | 2 | A49457 |
| 15 | 331 | 11.4 | 589 | 2 | T09059 |
| 16 | 330.5 | 11.3 | 1712 | 2 | T43210 |
| 17 | 324.5 | 11.1 | 689 | 2 | A38261 |
| 18 | 324.5 | 11.1 | 712 | 2 | T42960 |
| 19 | 322.5 | 11.1 | 1394 | 2 | A35626 |
| 20 | 321 | 11.0 | 2371 | 2 | S78549 |
| 21 | 314.5 | 10.8 | 3567 | 2 | T34513 |
| 22 | 313 | 10.7 | 798 | 2 | T22793 |
| 23 | 312.5 | 10.7 | 1820 | 2 | A55494 |
| 24 | 311 | 10.7 | 601 | 2 | B36346 |
| 25 | 309 | 10.6 | 683 | 2 | C36346 |
| 26 | 308 | 10.6 | 705 | 2 | S78040 |
| 27 | 303.5 | 10.4 | 1251 | 2 | S34968 |
| 28 | 303.5 | 10.4 | 1251 | 2 | A57293 |
| 29 | 297 | 10.2 | 2531 | 2 | A46019 |

| | | | | | | |
|----|-------|-----|------|---|--------|----------------------|
| 30 | 287 | 9.8 | 2471 | 2 | A49128 | cell-fate determin |
| 31 | 285 | 9.8 | 2352 | 2 | T30201 | Notch homolog prot |
| 32 | 284.5 | 9.8 | 2437 | 2 | S42612 | Transmembrane prot |
| 33 | 280.5 | 9.6 | 2703 | 1 | A24420 | Notch 3 protein - tr |
| 34 | 280 | 9.6 | 2318 | 2 | S45306 | extracellular prot |
| 35 | 279.5 | 9.6 | 387 | 2 | T38449 | Notch protein homo |
| 36 | 278 | 9.5 | 2531 | 2 | S18188 | hypothetical prote |
| 37 | 277.5 | 9.5 | 511 | 2 | T17298 | epidermal growth f |
| 38 | 276.5 | 9.5 | 493 | 2 | JC5621 | Notch protein homo |
| 39 | 273 | 9.4 | 2535 | 2 | A40043 | epidermal growth f |
| 40 | 269.5 | 9.2 | 1217 | 1 | ECMSMG | Notch B protein - |
| 41 | 264 | 9.1 | 1203 | 2 | A49175 | Notch homolog - se |
| 42 | 263.5 | 9.0 | 2531 | 2 | T11070 | Notch protein - Af |
| 43 | 256.5 | 8.8 | 2524 | 2 | A35844 | growth potentialin |
| 44 | 254.5 | 8.7 | 674 | 2 | T5476 | growth arrest-spec |
| 45 | 249 | 8.5 | 678 | 2 | B48089 | |

ALIGNMENTS

RESULT 1

THHUB

thrombomodulin precursor [validated] - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1988 #sequence_revision 12-May-1995 #text_change 15-Sep-2000

C:Accession: A41442; A28307; A29680; A27073; JX0264; S38954

R:Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Degashiki, Y.; Maru

J. Biochem. 103:281-285, 1988

A:Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed

A:Reference number: A41442; MUID:86227901

A:Accession: A41442

A:Molecule type: DNA

A:Residues: 1-575 <SHI>

A:Cross-references: DDBJ:D00210; NID:g220126; PIDN:BA00149.1; PID:g220127

R:Jackman, R.W.; Beeler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.

Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987

A:Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of th

A:Reference number: A28307; MUID:87317665

A:Accession: A28307

A:Molecule type: DNA; mRNA

A:Residues: 1-472, 'A', 474-575 <JAC>

A:Cross-references: GB:J02973; NID:g339658; PIDN:AAA61175.1; PID:g339659

R:Suzuki, K.; Kusumoto, H.; Degashiki, Y.; Nishioke, J.; Maryama, I.; Zushl, M.;

EMBO J. 6, 1891-1897, 1987

A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on

A:Reference number: A29680; MUID:88004395

A:Accession: A29680

A:Molecule type: mRNA

A:Residues: 1-575 <SUZ>

A:Cross-references: GB:X05495; NID:g37123; PIDN:CAA29045.1; PID:g736251

A:Experimental source: Lung endothelium

A:Note: part of this sequence, including the amino end of the mature protein, were

R:Men, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.

Biochemistry 26, 4350-4357, 1987

A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization

A:Reference number: A27073; MUID:88024950

A:Accession: A27073

A:Molecule type: mRNA

A:Residues: 1-472, 'A', 474-575 <MEN>

A:Cross-references: GB:M15552; NID:g339656; PIDN:AB59508.1; PID:g339657

A:Experimental source: Placenta

A:Note: parts of this sequence were determined by protein sequencing

R:Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Okuchi, M.; Kimura, S.; Aoki, N.

J. Biochem. 113, 433-440, 1993

A:Title: Urinary thrombomodulin, its isolation and characterization.

A:Reference number: JX0264; MUID:93293792

A:Accession: JX0264

A:Molecule type: protein; mRNA

A:Residues: 19-472, 'A', 474-486 <YAM>

A:Experimental source: urine

A:Note: the urinary form appears to be identical with that circulating in plasma

R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell,

Biochem. J. 295, 131-140, 1993
 A:Title: Identification of the predominant glycosaminoglycan-attachment site in soluble serine.
 A:Reference number: S38954; MUID:94029900
 A:Accession: S38954
 A:Molecule type: protein
 A:Residues: 475-491, 'X', 493-494 <GER>
 A:Note: the residue designated 'X' was determined to be a Ser with covalently bound chori R.Melinger, D.P.; Komives, E.A.
 A:Reference number: A67369; PDB:1ZAO
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue R.Melinger, A.; Mathews, I.T.
 A:Reference number: A32804; PDB:1HJF
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442
 R.Hirshal, R.; Komives, E.A.; M., F.
 submitted to the Brookhaven Protein Data Bank, November 1995
 A:Reference number: A65583; PDB:1FGD
 A:Contents: annotation; conformation by (1)H-NMR, residues 427-444
 R.Hirshal, R.; Komives, E.A.; M., F.
 Protein Sci. 5, 195-203, 1996
 A:Title: Structural resiliency of an EGF-like subdomain bound to its target protein, the A:Reference number: A58595; MUID:96276211
 A:Contents: annotation; conformation by (1)H-NMR
 C:Genetics:
 A:Gene: GDB:THBD
 A:Cross-references: GDB:119613; OMIM:188040
 A:Map position: 20p11.2-20p11.2
 A:Introns: #status absent
 C:Complex: homodimer, urinary form
 C:Function:
 A:Description: inhibits thrombin activation of fibrinogen; cofactor for thrombin activation
 A:Pathway: blood coagulation/moderation
 A:Note: the membrane-bound form is located on the endothelium luminal surface of arterie A:Note: thrombin complexed with the membrane-bound form is subject to endocytosis
 C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
 C:Keywords: anticoagulant; beta-hydroxyasparagine; beta-hydroxyaspartic acid; blood coag e protein
 E:1-18/Domain: signal sequence #status predicted <SIG>
 E:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>
 F:19-513/Domain: extracellular #status predicted <EXT>
 F:19-486/Product: thrombomodulin, urinary form #status experimental <MAU>
 F:24-167/Domain: C-type lectin homology <LCH>
 F:177-199/Region: PEST sequence
 F:201-233/Region: PEST sequence
 F:245-280/Domain: EGF homology <EG1>
 F:288-323/Domain: EGF homology <EG2>
 F:329-362/Domain: EGF homology <EG3>
 F:369-404/Domain: EGF homology <EG4>
 F:408-439/Domain: EGF homology <EG5>
 F:445-480/Domain: EGF homology <EG6>
 F:485-513/Region: PEST sequence
 F:517-539/Domain: transmembrane #status predicted <TMN>
 F:540-575/Domain: intracellular #status predicted <INT>
 F:47,115,116,382,409/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:174,223,411,504/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F:245-256,265,267-280,288-296,292-308,310-323,329-340,336-349,351-362,369-378,374-38 F:334,498/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F:342/Modified site: erythro-beta-hydroxyasparagine (Asn) #status experimental
 F:490,492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimental

Query Match 99.9%; Score 2912; DB 1; Length 575;
 Best local similarity 99.8%; Pred. No. 3e-166;
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVIVLALAGLGGPAPAPQPGSGQVHCDFALVPGPTFLNASQICDGRGHM 60
 DB 1 MGVIVLALAGLGGPAPAPQPGSGQVHCDFALVPGPTFLNASQICDGRGHM 60
 QY 61 TVRSSVAADVISLLNGDGVGRRRLWIGLDPGCGDKRLGPRFGFWTGDNNTSYS 120
 DB 1 TVRSSVAADVISLLNGDGVGRRRLWIGLDPGCGDKRLGPRFGFWTGDNNTSYS 120

DB 61 TVRSSVAADVISLLNGDGVGRRRLWIGLDPGCGDKRLGPRFGFWTGDNNTSYS 120
 QY 121 RMRRLDNLNAPLCGPACVAVSAEATVSEPTINEQCEYKACGFCFEPFATCPPLAV 180
 DB 121 RMRRLDNLNAPLCGPACVAVSAEATVSEPTINEQCEYKACGFCFEPFATCPPLAV 180
 QY 181 EPGAAAASVITYGTPPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVQGNMAREAP 240
 DB 181 EPGAAAASVITYGTPPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVQGNMAREAP 240
 QY 241 GAMDGVENGSGEHCNMPGAPRCQCPAGAAIADGRSCAASCTGSCNLCGHPVNP 300
 DB 241 GAMDGVENGSGEHCNMPGAPRCQCPAGAAIADGRSCAASCTGSCNLCGHPVNP 300
 QY 301 DPGSYSCMCEGYRLADQHRCEVDDCILEPSPQRCVNTGGEFCHYPNYDLVDG 360
 DB 301 DPGSYSCMCEGYRLADQHRCEVDDCILEPSPQRCVNTGGEFCHYPNYDLVDG 360
 QY 361 ECEPVPDCEFRANCEYCCQPLNQTSTYLCVCAEGFAPLPHEPHRQMCNCTACPADCDN 420
 DB 361 ECEPVPDCEFRANCEYCCQPLNQTSTYLCVCAEGFAPLPHEPHRQMCNCTACPADCDN 420
 QY 421 TQASCCECPGYILDDGFICTDIDECENGFCGVCVHMLPGTFECICGSDALRHITGDC 480
 DB 421 TQASCCECPGYILDDGFICTDIDECENGFCGVCVHMLPGTFECICGSDALRHITGDC 480
 QY 481 DSGKVDGSDSGSEPPSPPTPTLPPAVGLVHSG 516
 DB 481 DSGKVDGSDSGSEPPSPPTPTLPPAVGLVHSG 516

RESULT 2
 A60501
 thrombomodulin precursor - mouse
 N:Alternate names: fetomodulin
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
 C:Accession: S08488; A32001; A60501
 R:Dittman, W.A.; Majerus, P.W.
 Nucleic Acids Res. 17, 802, 1989
 A:Title: Sequence of a cDNA for mouse thrombomodulin and comparison of the predicted A:Reference number: S08488; MUID:89128454
 A:Accession: S08488
 A:Molecule type: mRNA
 A:Residues: 1-577 <DIR>
 A:Cross-references: EMBL:X14432; NID:954781; PIDD:CA32597.1; PID:954782
 R:Dittman, W.A.; Kumada, T.; Sadler, J.E.; Majerus, P.W.
 J. Biol. Chem. 263, 15815-15822, 1988
 A:Title: The structure and function of mouse thrombomodulin. Phorbol myristate acetat A:Reference number: A32001; MUID:89008498
 A:Accession: A32001
 A:Molecule type: mRNA
 A:Residues: 97-577 <DIR>
 A:Cross-references: GB:J04060
 R:Imada, S.; Yamaguchi, H.; Nagumo, M.; Katayanagi, S.; Iwasaki, H.; Imada, M.
 Dev. Biol. 140, 113-122, 1990
 A:Title: Identification of fetomodulin, a surface marker protein of fetal development A:Reference number: A60501; MUID:90292331
 A:Accession: A60501
 A:Molecule type: protein
 A:Residues: 19-22;330-343;479-489;545-555;562-575 <IMN>
 C:Comment: Thrombomodulin binds to and internalizes with thrombin. It is also a cofac C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
 C:Keywords: anticoagulant; membrane protein; phosphoprotein; receptor
 F:24-165/Domain: C-type lectin homology <LCH>
 F:244-279/Domain: EGF homology <EG1>
 F:287-322/Domain: EGF homology <EG2>
 F:328-361/Domain: EGF homology <EG3>
 F:368-403/Domain: EGF homology <EG4>
 F:407-438/Domain: EGF homology <EG5>
 F:444-479/Domain: EGF homology <EG6>

Wed May 8 16:19:32 2002

us-09-509-994-2.rpt

Page 3

Query Match 65.3%; Score 1903; DB 2; Length 577;
Best Local Similarity 65.6%; Pred. No. 2,9e-106;
Matches 342; Conservative 46; Mismatches 125; Indels 8; Gaps 4;

QY 1 MGVTVLGLATAGLGFAPAPAPPOGSGCVENHDCPLALYEPAPFLUASQICDGLGHLM 60
DB 1 MGVTVLGLATAGLGFAPAPAPPOGSGCVENHDCPLALYEPAPFLUASQICDGLGHLM 60
QY 61 TVRSVVAADVLSLLNGDG-GVGRRLMTIGLQRPCCGDPKRLPLRGFWMTGDNNTSY 119
DB 61 TVRSVVAADVLSLLNGDG-GVGRRLMTIGLQRPCCGDPKRLPLRGFWMTGDNNTSY 119
QY 120 SRNARLDLNGAPLCPGLCYAASAAATVPSEPIPEEQCEKADGFLCEHPFATGCPPLA 179
DB 118 SRNARLDLNGAPLCPGLCYAASAAATVPSEPIPEEQCEKADGFLCEHPFATGCPPLA 177
QY 180 VEP-GAAAAAVSTTGTTPPAAGADPQALPVGSSAAVAPLGLQMTAPPGAVOGHARE 238
DB 178 VNTRDPEAAHISTVNTPTPGVSGADEPQLPVGSSAAVAPLGLQMTAPPGVSGHAWME 237
QY 239 APGAMDCSVENGCSEHACNAIPGAPRCOCPAGALOADGRCSTASATQSCNDLCEHCVY 298
DB 238 ATGAMNCSYVENGCSEHACNAIPGAPRCOCPAGALOADGRCSTASATQSCNDLCEHCVY 297
QY 299 NPDPGSGYSCMETGTGRLAADHRCEDVDCLLEPSPQRCVNTGGFECHECPANVDY 358
DB 298 NAEVPGSYSCMETGTGRLAADHRCEDVDCLLEPSPQRCVNTGGFECHECPANVDY 357
QY 359 DGECEVDPDFCFRANCEYOCPLNOSTYLVCACBEGFAPLPHPRHCOMFQNTCAPADCD 418
DB 358 DGECEVDPDFCFRANCEYOCPLNOSTYLVCACBEGFAPLPHPRHCOMFQNTCAPADCD 417
QY 419 PNTQASCECEGEGYITLDGFCITDIDCENGSCGVCNLTGTEGICGPPSALARIHGT 478
DB 418 PNTQASCECEGEGYITLDGFCITDIDCENGSCGVCNLTGTEGICGPPSALARIHGT 477
QY 479 DCDGSKV--DGDGSGSGEPSPPTPGSTLPPAVGLVHSG 516
DB 478 DCDPIVREDTKEEGSGEPSPPTPGSTLPPAVGLVHSG 518

RESULT 3

A25918
Chromodomain - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: A25918
R:Jackson, R.W.; Beeler, D.L.; Vandewater, L.; Rosenberg, R.D.
Proc. Natl. Acad. Sci. U.S.A. 83, 8834-8838, 1986
A:Title: Characterization of a chromodomain cDNA reveals structural similarity to the
A:Reference number: A25918; MUID:87067408
A:Accession: A25918
A:Molecule type: mRNA
A:Residues: 1-356 <JAC>
A:Cross-references: GB:M14657; NID:9163762; PID:AAA30785.1; PID:9163763
C:Superfamily: chromodomain; C-type lectin homology; EGF homology
C:Keywords: transmembrane protein
F:21-56/Domain: EGF homology <EG3>
F:64-97/Domain: EGF homology <EG3>
F:103-116/Domain: EGF homology <EG3>
F:143-178/Domain: EGF homology <EG4>
F:182-213/Domain: EGF homology <EG5>
F:219-253/Domain: EGF homology <EG6>

Query Match 37.6%; Score 1095.5; DB 2; Length 356;
Best Local Similarity 63.4%; Pred. No. 2e-58;
Matches 189; Conservative 31; Mismatches 63; Indels 15; Gaps 6;

QY 229 CAVGCHNAREAPGAMDCSVENGCSEHACNAIPGAPRCOCPAGALOADGRCSTASATQSC 288
DB 229 CAVGCHNAREAPGAMDCSVENGCSEHACNAIPGAPRCOCPAGALOADGRCSTASATQSC 288
QY 5 GETEGRSREAPGAMACVERGGCGHCKSGSAGASNCILCPADALOADGRCSTASATQSC 64
DB 5 GETEGRSREAPGAMACVERGGCGHCKSGSAGASNCILCPADALOADGRCSTASATQSC 64

QY 289 NDLCSEHCVPNPDGSGYSCMETGTGRLAADHRCEDVDCLLEPSPQRCVNTGGF 348
DB 65 HOLCEHVC--HLGLGNTYICIEAGYOLAADHRCEDVDCLLEPSPQRCVNTGGF 122
QY 349 CHCVPNVDYDGECEVDPDFCFRANCEYOCPLNOSTYLVCACBEGFAPLPHPRHCOMFC 408
DB 123 CHCVPNVDYDGECEVDPDFCFRANCEYOCPLNOSTYLVCACBEGFAPLPHPRHCOMFC 182
QY 409 NQTPADCDPNTQASCECEGEGYITLDGFCITDIDCENGSCGVCNLTGTEGICGPPSALARIHGT 468
DB 183 NQTPADCDPNTQASCECEGEGYITLDGFCITDIDCENGSCGVCNLTGTEGICGPPSALARIHGT 241
QY 469 DSALARIHGTDCSGK-----DGDGSGSGEPSPPTPGSTLPPAVGLVHSG 516
DB 242 DSALARIHGTDCSGK-----DGDGSGSGEPSPPTPGSTLPPAVGLVHSG 297

RESULT 4

T13954
MEGF6 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T13954
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like mo
A:Reference number: Z14126; MUID:98360089
A:Accession: T13954
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1574 <NAB>
A:Cross-references: EMBL:AB011532; NID:93449293; PID:BA42462.1; PID:93449294
A:Experimental source: strain Sprague-Dawley; brain
A:Gene: MEGF6

Query Match 14.2%; Score 414; DB 2; Length 1574;
Best Local Similarity 34.1%; Pred. No. 2.2e-17;
Matches 104; Conservative 28; Mismatches 101; Indels 72; Gaps 17;

QY 224 CTAPPG-AVOGHNAREAPGAMDCSVENGCSEHACNAIPGAPRCOCPAGALOADGRCSTASATQSC 282
DB 147 CRCPGTQLOD-CKTQDVIDECRANGCGQRCVNTGSGYSCMETGTGRLAADHRCEDVDCLLEPSPQRCVNTGGF 204
QY 283 SATQSC--NDLCSEHCV-----NPDPGSGYSCMETGTGRLAADHRCEDVDCLLEPSPQRCVNTGGF 309
DB 205 AISCSTLGNCGCOHCVOQLTVQHRCCRPQYOLOEDGRRCVRSPCABGNGCMHIQ 263
QY 310 -----CEGTGRLAADHRCEDVDCLLEPSPQRCVNTGGFCHCVPNVDYD 359
DB 264 ELRGALHCGHPRGYOLAADHRCEDVDCLLEPSPQRCVNTGGFCHCVPNVDYD 323
QY 360 G-EC---VEVPDPCFRAN--CEYCOPLNOSTYLVCACBEGFAPLPHPRHCOMFC--- 408
DB 324 GRCQRIEMETVNSCEANGNGCSHCSH-TSTGPICTCPRGY-----ELDEDQKTCIDID 377
QY 409 ---NOTACPADCDPNTQASCECEGEGYITLDGFCITDIDCENGSCGVCNLTGTEGICGPPSALARIHGT 461
DB 378 DCANSPCCQANC-ANTPGYEGSCGFAGYRLNTDGGCCEDVDCAAGHGSGEHHGSLAGS 436
QY 462 FECLC 466
DB 437 FQCEC 441

RESULT 5

T27283
Hypothetical protein Y6AG10A.f - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27283
R:Ainscough, R.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20356
A:Accession: Y27283
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1620 <WIL>
A:Cross-References: EMBL:AL10498; NID:61542303; PIDN:GAB54471.1; CESP:Y64G10A.f
A:Experimental source: clone Y64G10A
C:Genetics:
A:Gene: CESP:Y64G10A.f
A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1; 7

Query Match 14.2%; Score 414; DB 2; Length 1620;
Best Local Similarity 27.0%; Pred. No. 2.2e-17;
Matches 128; Conservative 38; Mismatches 176; Indels 132; Gaps 20;

QY 119 YSRMARLDNGAPLGPICVAVSAE-----ATPSEPIWEEOCE--YVADGCEEF 169
DB 56 YLRPARFRRSGSKCCLLRQANGSADLCNNGTGVPSHNDNEQVCEPVGFTAKQY 115
QY 170 HEPATGRPLAVPPGAAAASVITYGTFPAARGADFOALPYGSSAAVAPLGLMCTAPPG 229
DB 116 D-ANEC--MANNGCEHCN-TIGTY-----CRCMPG 145
QY 230 AVQGHAREAPGAMDCSEVNGCEHCNNAIPGAPRCQCPAGALQADRSCTASATQSC- 288
DB 146 FELSGDNATCSDIDCAVNSGCGSDRCVNSPGFRCDCPSDLIYHADRTG--GKVTSCS 203
QY 289 --NDLCHEHCYVNPQPSYSCMCEMYRIADQHRCEVDCLILEPSPCPQRCVNTQGG 346
DB 204 TDNGGCEHC--ENDSNGEFTYRCRVFKLSEKRSQYVDFPDNKGCGCHCHNNHGR 262
QY 347 FECHCYPNYL-VDECEVYDPCFRAN-CEYOCQPLNOTSYLCVACGAFAPRPHPRC 404
DB 263 ACCOCYPGFHLSDYRSCVDIDECANNGCEHCENVKGT-YKCKREGY-OLGDSGRYC 320
QY 405 OME-----CNOTACPADC--DPNTQASCEPEGYL----- 433
DB 321 EEMLGCGQYNGCGQHDYDQPDGHHCKCRNGYILANQKICHONISVTHARAPRLMD 380
QY 434 -----DGF-----ICTDDEC-ENKGCSCGYC 455
DB 381 STEYTCVTPDLCYCHKMLDSGHVQCFCDDGYELIDSKRFQDIDNECHENNGCSDIC 440
QY 456 HMLPTFECICGDSALARAIGT-----DCDS-----GKVDGSGSGSEP 495
DB 441 VNLASGVCECKPGRFLMKDRKTCEDISECSSNNGCEOICSNQEGYKSCSEP 494

RESULT 6
Fibrillin-2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Feb-1996 #sequence, revision 23-Feb-1996 #text_change 11-Jan-2000
C:Accession: A57278
R:Zhang, H.; Hu, W.; Ramirez, F.
J. Cell Biol. 129, 1165-1176, 1995
A:Title: Developmental expression of fibrillin genes suggests heterogeneity of extracellular matrix
A:Reference number: A57278; MUID:95263670
A:Accession: A57278
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2907 <ZHA>
A:Cross-References: GB:139790; NID:9762830; PIDN:AAA74908.1; PID:9762831
F:1239-1274/Domain: EGF homology <EGF1>
F:2488-2523/Domain: EGF homology <EGF>

Query Match 12.8%; Score 374; DB 2; Length 2907;
Best Local Similarity 32.9%; Pred. No. 8.5e-15;
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

QY 244 DCSVENGCCEHCNNAIPGAPRCQCPAGALQADRSCTASATQSCNDLCHEFCVPPND-- 301
DB 1238 EGMIMNGCCTOCTNSGSEYSCSEGVYALMDPGRSCA-----DIDE--CENNPDIC 1287
QY 302 -----QPSYSCMCEYRIADQHRCEVDCLILEPSPC-QRCVNTQGGFECHCP 353
DB 1288 DGGCCTNIPGEXKRLCYDGFMAEMDKITCVNEDCLNPICLSGTENTKSGFICHCDM 1347
QY 354 NYDLVDE--CPEVPDPC--FRANCEYOCQPLN-OTSYLCVACGFA-----PIP 398
DB 1348 GYSVAKGKTGCTD-VDECEIGHANCDMHASCLNVPSGSCSRGHWNGIKICIDIDECA 1406
QY 399 HEPHRCOMFCNOTACPADCDPNTQAS--CEPEGYILDDGFTCYDIDE-----CENG- 448
DB 1407 NGTHQCSF-----NAC-VNTPGSRACSEGF-TGDGFTCSVDDECAENNLCEMO 1457
QY 449 -----GF-----GCHNLPTGFECICGDSALA 473
DB 1458 CLNPGAYRCCEKGFPPASDSRSCQDIDECSTFONICVFGICNNLPMFHCIDCDGYELD 1517
QY 474 RHIG--TDCCD 481
DB 1518 RTGNCCTDID 1527

RESULT 7
Fibrillin 1 - bovine
A:Accession: A55567
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Feb-1995 #sequence, revision 20-Feb-1995 #text_change 11-Jan-2000
C:Accession: A55567
R:Willsta, D.; J., L.; Potter, K.A.; Womack, J.; Byers, P.H.
Genomics 23, 480-485, 1994
A:Title: Sequence of the coding region of the bovine fibrillin cDNA and localization
A:Reference number: A55567; MUID:95137597
A:Accession: A55567
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2871 <WIL>
A:Cross-References: GB:128748; NID:9508427; PIDN:AAA74122.1; PID:9508428
F:Superfamily: unassigned EGF-related proteins; EGF homology
F:1201-1236/Domain: EGF homology <EGF>

Query Match 12.6%; Score 368; DB 2; Length 2871;
Best Local Similarity 31.5%; Pred. No. 1.9e-14;
Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps 20;

QY 244 DCSVENGCCEHCNNAIPGAPRCQCPAGALQADRSCTASATQSCNDLCHEFCVPPND-- 300
DB 1200 EOSTMNGCCTCTNSGSEYSCSEGVYALMDPGRSCA-----DIDE--CENNPDIC 1249
QY 301 -----DQPSYSCMCEYRIADQHRCEVDCLILEPSPC-QRCVNTQGGFECHCP 353
DB 1250 DGGCCTNIPGEXKRLCYDGFMAEMDKITCVNEDCLNPICLSGTENTKSGFICHCDM 1309
QY 354 NYDLVDE--CPEVPDPC--FRANCEYOCQPLN-OTSYLCVACGFA-----PIP 398
DB 1310 GYSKRRKKTGCTD-INCCEIGHANCDMHASCLNVPSGSCSRGHWNGIKICIDIDECA 1406
QY 399 HEPHRCOMFCNOTACPADCDPNTQAS--CEPEGYILDDGFTCYDIDE-----CENG- 448
DB 1369 NGTHQCSF-----NAC-VNTPGSRACSEGF-TGDGFTCSVDDECAENNLCEMO 1457
QY 449 -----GF-----GCHNLPTGFECICGDSALA 473
DB 1420 CLNPGAYRCCEKGFPPASDSRSCQDIDECSTFONICVFGICNNLPMFHCIDCDGYELD 1517
QY 474 RHIG--TDCCD 481
DB 1480 RSGGCTDVNECDPTTCTISGNCVNTPGSYTCDCCPD---FELNPRVGCV 1527

RESULT 8
A55184
Fibulin-2 precursor - human
N:Alternate names: protein DKFZP586A1519.1
C:Species: Homo sapiens (man)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
C:Accession: A55184; T08744
R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.
GenBank, 22, 425-430, 1994
A:Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the gene
A:Reference number: A55184; MUID:95104855
A:Accession: A55184
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1184 <ZHA>
A:Cross-references: GB:X82494; NID:9575232; PID:CA57876.1; PID:9575233
R:Mambuti, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z16471
A:Accession: T08744
A:Molecule type: mRNA
A:Residues: 656-719, ODECLMGADHCSRROFCVNTLGSFYCVNHTVLCADGYILNAHRKCYD, 720-853, 'T', 855-1111
A:Cross-references: EMBL:AL050095
A:Experimental source: adult uterus; clone DKFZP586A1519
C:Genetics:
A:Gene: GDB:FBLN2
A:Cross-references: GDB:293037; OMIM:135821
A:Map position: 3p25-3p24
A:Note: DKFZP586A1519.1
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing; extracellular matrix
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-1184/Product: fibulin-2 protein #status predicted <MAT>
F:905-941/Domain: EGF homology <EGF>

Query Match 12.6%; Score 367.5; DB 2; Length 1184;
Best Local Similarity 30.0%; Pred. No. 1e-14;
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;
154 EEOQCEVAD-GFLCEFFPATC-----RPLAV-----EGGAAAVSTTGPFAARG 201
Db 531 EGQSCSEPNLGLPCN-HVMLSCGSEBPLVPEYRRPPEAARPRYS-----EEM 582
OY 202 ADFQALPGSSAAV---APLGQLMCTAPPGAYOGHMAKAPGAMDCSEVNGCEHACNA 258
Db 563 AGKALSLGTENELPNSLPDGDDECLLPLSEL-----COHLICIN 622
OY 259 IPGADRCOCPPAGALQADGRSC-----TASAT-----QS 287
Db 623 TVGSYHACFPGLSLQDDGRTRCPREGRHPPPEAPQEPALKEFSQVANSNTIPLPLPONT 682
OY 288 CND--LCHEFCYRPNDDGSGYSCKMETGYRLAADOHREDVDCLLESPCP--QRQVNT 343
Db 683 CNDNCPCKQVC---STVGSALCSCFPGYALMADGVSEDIINECVTDLHTSCRGEHCYNT 739
OY 344 OGGEFCH---CYPNYDLVDECEVPVPCFRANCEYOCQPLNTSYLCV-----389
Db 740 LGSFHCYKALCEPGLYALKDEBC-EDVDEC--AMGTHICQF---GLFCQNTKSFYCOA 792
OY 390 ---CAEGFAPLPH-----EPHRCQMFQNTCAPDADCPNTQASCECPGY-I 432
Db 793 RORCMGFLQDPGEGNCVDINETSLSGPCRPGFSCINTVGSYTCORNFLL---CARGYHA 849
OY 433 LDDGFLCTDIDECENGFCSG---VCHNLPGTFEFCIC 466
Db 850 SDDGAKCYDVNCEYGVHRCGEGOVCHNLPGSYRCDC 886

C:Species: Homo sapiens (man)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 17-Nov-2000
C:Accession: A54105; S17063; S31101
R:Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sanguinetti, C.; Bonadio, J.; Mecha
J. Cell Biol. 124, 855-863, 1994
A:Title: Structure and expression of fibritillin-2, a novel microfibrillar component pr
A:Reference number: A54105; MUID:94165150
A:Accession: A54105
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-2918 <ZHA>
A:Cross-references: GB:U032272
R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras,
Nature 352, 330-334, 1991
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff
A:Reference number: S17063; MUID:91304567
A:Accession: S17063
A:Molecule type: mRNA
A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLT', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>
A:Cross-references: EMBL:X62009
R:Milwicz, D.M.
submitted to the EMBL Data Library, December 1992
A:Reference number: S31101
A:Accession: S31101
A:Molecule type: mRNA
A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLT', 1922-1923, 'LD', 1926, 'P',
A:Cross-references: EMBL:X62009
C:Genetics:
A:Gene: GDB:FBN2
A:Cross-references: GDB:128122; OMIM:121050
A:Map position: 5q23-5q31
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: extracellular protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-2918/Product: fibritillin-2 #status predicted <MAT>
F:1245-1280/Domain: EGF homology <EGF1>
F:1970-2013/Domain: EGF homology <EGF2>

Query Match 12.6%; Score 366; DB 2; Length 2918;
Best Local Similarity 33.1%; Pred. No. 2.5e-14;
Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps 17;
OY 244 DCSVENGCSEHACNAIPGRCQPPAGALQADGRSCASATQSCNDLCEHFCVNPDP-- 301
Db 1244 ECMIMNGGCTQGTNTEGSEGCSEGYALMPGRCA-----DIDE--CENNPDC 1293.
OY 302 -----QPGSYSCKMETGYRLAADOHRCEDVDCLLESPCP--QRQVNTQGFECYCP 353
Db 1294 DGGQCTNIPGKYRCLCYDGFMSMDKKTIDVNECLNINICMFGECEYTKSFTCHQL 1353
OY 354 NYDLVGE--CYEVPVPC--FRANCEYOCQPLN-QTSYLCVCAEGFAPLPH-----EPHR 403
Db 1354 GYVKKGTCTCTD-VDECELGANNCMHASCLNIPGSFSCSCEH--IGNGIKCIDLDE 1410
OY 404 CQFNCQNTACPADCPNTQAS--CECPGYILDGFCITDIDE-----CENG-----448
Db 1411 CSNGTHQCSINAGC-VNTPGSYRCACSEGF-ITDGTGSDVDVBCAENMLCENGQCLNVP 1468
OY 449 -----GF-----CSGYCHNLPGTFEFCICPDSALAHNG- 477
Db 1469 GAYRCCEMGFTPADSDRSQDIDECSEFQNICVSGTCNNLPGMFHCICDDGYELDRGTGN 1528
OY 478 -TDGD 481
Db 1529 CTDID 1533

RESULT 9
A54105
Fibritillin-2 precursor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 02-Jun-1995 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000

C:Accession: A47221; I54355; S17064; I59574; S17062; S62111; A34198
 R:Corson, G.M.; Chaberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.
 Genomics 17, 476-484, 1993
 A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure
 A:Reference number: A47221; MUID:94010947
 A:Accession: A47221
 A:Molecule type: mRNA
 A:Residues: 1-337, 'T', 339-1029 <COR>
 A:Cross-references: GB:X63556
 R:Perle, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Panglman, T.; Bonad
 Hum. Mol. Genet. 2, 961-968, 1993
 A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene F
 A:Reference number: I54355; MUID:93372860
 A:Accession: I54355
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 133-3002 <PER>
 A:Cross-references: GB:U3923; NID:9306745; PIDN:ABR02036.1; PID:9306746
 R:Malet, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.
 Nature 352, 334-337, 1991
 A:Title: Partial sequence of a candidate gene for the Marfan syndrome.
 A:Reference number: S17064; MUID:91304568
 A:Accession: S17064
 A:Molecule type: mRNA
 A:Residues: 1030-3002 <MAS>
 A:Cross-references: EMBL:X63556
 R:Dietz, H.C.; Valle, D.; Franccomano, C.A.; Kendzior, R.J.
 Science 259, 680-683, 1993
 A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.
 A:Reference number: I59574; MUID:93157831
 A:Accession: I59574
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 2217-2288, 'T', 2290-2325 <RES>
 A:Cross-references: GB:S54426; NID:9264860; PIDN:ABR2544.1; PID:9264861
 R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P.
 Nature 352, 330-334, 1991
 A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differe
 A:Reference number: S17062; MUID:91304567
 A:Accession: S17062
 A:Molecule type: mRNA
 A:Residues: 'VUTVVFITLSTYKML', 944-1444 <LEI>
 A:Cross-references: EMBL:X62008; NID:931398; PIDN:CA56534.1; PID:95294015
 A:Accession: S62111
 A:Molecule type: Protein
 A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LE2>
 R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.
 J. Biol. Chem. 264, 21381-21385, 1989
 A:Title: Connective tissue microfibrils. Isolation and characterization of three large f
 A:Reference number: A34198; MUID:90078246
 A:Accession: A34198
 A:Molecule type: Protein
 A:Residues: 565-575; 1890-1892, 'T', 1894-1900 <MAD>
 C:Comment: Fibrillin is a major component of elastin-associated microfibrils.
 C:Genetics:
 A:Gene: GDB:FBNI
 A:Cross-references: GDB:127115; OMIM:134797; OMIM:154700
 A:Map position: 15q21.1-15q21.1
 A:Introns: 2236/1; 2258/1; 2297/1
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein; M
 F:1-3002/Product: fibrillin (5'-region exon A splice form) (fragment) *status predicted
 F:1332-3002/Product: fibrillin (5'-region exon C splice form) *status predicted <MATIC>
 F:1332-1367/Product: EGF homology <EGF>
 F:1457-1492/Product: EGF homology <EGF2>
 F:2262-2295/Product: EGF homology <EGF1>

Query Match 12.5%; Score 364; DB 2; Length 3002;
 Best Local Similarity 30.8%; Pred. No. 3.4e-14;
 Matches 11; Conservative 22; Mismatches 105; Indels 122; Gaps 20;

QY 244 DCSVNGGCEHACNAIFGAPRCQCPAGALQADGNSCTASATQSCNDLCEHFCVNP--- 300

Db 1331 ECSTIMNGCTFTNBSGSEYSCQDPGFALMPQORST-----DIDE--CEDPNIC 1380
 QY 301 -----DQPSYSCMETGYRLAADHCEEDVDCTLEPSPC-PORCVNTQGGFECYCP 353
 Db 1381 DGGCTNIPGEYRCLCTGDMASEDMKTCYDVNECDLNINICLSTGCEINTKSFICHO 1440
 QY 354 NYDLVDGE-CVEPVDP-C-FRANCYQOPLNQT-SYLCVAEGFA-----PIP 398
 Db 1441 GYSOKKKTGCTD-INCEIGAHNCKHAYCTNAGSFKSCSPGWIIGIKCTDLECS 1499
 QY 399 HEPRHRCMFCNQACAPDCDPTNQA--CEPCGYTLIDGFTCTDIDE--EN----- 447
 Db 1500 NGTHMCSQH-----ADC-KNTMGSTRCLCKEGY-TSGFTCTDIDESENLNLCGNG 1550
 QY 448 -----GGF-----CS-----GVCHNLPGEFICGPDPSALA 473
 Db 1551 CUNAPGTYRCECDMGFPYPSADGKACEDIDECSLPNICVGTGNNLGLRCEGEIYELD 1610
 QY 474 RHIG-----TDCSGKVDGDSGSGEPSPPTGCS-----TLTPPAGLV 513
 Db 1611 RSGGNCYDVNECDLPTTICISGNCVN-----TSGSYICDPCPDPELMPFRVGCV 1658
 RESULT 11
 A35624
 Fibrillin-1 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 11-Jan-2000
 C:Accession: A55624
 R:Yin, W.; Smiley, E.; Gemmiller, J.; Sanglincet, C.; Lawton, T.; Perle, L.; Ramire
 J. Biol. Chem. 270, 1798-1806, 1995
 A:Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin
 A:Reference number: A55624; MUID:95130561
 A:Accession: A55624
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 12871 <YIN>
 A:Cross-references: GB:U29454; NID:9575509; PIDN:AAA56840.1; PID:9575510
 C:Genetics:
 A:Gene: Fbn-1
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 F:1201-1236/Product: EGF homology <EGF>

Query Match 12.4%; Score 362; DB 2; Length 2871;
 Best Local Similarity 29.4%; Pred. No. 4.3e-14;
 Matches 117; Conservative 26; Mismatches 127; Indels 128; Gaps 21;

QY 212 SAATAPVIGL-----QLMCTAPRGAVGHNAREAPGAMQSVNNGCEHACNAIFGAPRC 265
 Db 1162 SANLCFPHGRCVNLIGKYQACANPEYHPTHDLFLFCVDIDECSIMNGCEFTTNSDSTEC 1221
 QY 266 QCPAGALQADGRCSTASATQSCNDLCEHFCVNP-----DQPSYSCMETGYRL 316
 Db 1222 SCQGFALMPDQGRST-----DIDQ--CEBNPNICDGGCTNIPGEYRCLCYGFMA 1271
 QY 317 AADQHRCEVDVDCITLESPPC-PORCVNTQGGFECYCPNYDLVGE--CVEPVDP-C-FR 371
 Db 1272 SEDKKTQDVNEDCLNINICLSTGCEINTKSFICHOCHMGSGKRGKTGCTD-INCEIGA 1330
 QY 372 ANCYQOQPLNQT-SYLCVAEGFA-----PIPHPRHCQMFQNTACAPACDPN 420
 Db 1331 HNCRAHVCNTNAGSFKSCSPGWIIGIKCTDLECSNGTHMCSQH-----ADC-KN 1382
 QY 421 TQAS--CEPEGYTLIDGFTCTDIDE--EN-----GGF----- 450
 Db 1383 TMGSTRCLCKDGY-TSGFTCTDIDESENLNLCGNGCUNAGYRCECDMGFPYPSADG 1441
 QY 451 -----CS-----GVCHNLPGEFICGPDPSALRHIG-----TDCDSGK 484
 Db 1442 KACEDIDECSLPNICVGTGNNLGLRCEGEIYELDRSGGNCYDVNECDLPTTICISGN 1501

OY 485 VDGSDSGSGPPSPRPS-----TLTPPAGLV 513
 DB 1502 CVN-----TPGSTYCCDCSPDELPNTRVGCV 1527

RESULT 12

246488
 hypothetical protein DKFZp434J065.1 - human (fragment)

C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C:Accession: T46488
 R:Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: 223035
 A:Accession: T46488

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-741 <AAA>
 A:Cross-references: EMBL:AL137638
 A:Experimental source: adult testis; clone DKFZp434J065
 A:Note: DKFZp434J065.1

Query Match 12.2%; Score 357; DB 2; Length 741;
 Best Local Similarity 31.6%; Pred. No. 2, 9e-14;
 Matches 86; Conservative 40; Mismatches 102; Indels 44; Gaps 15;

OY 245 CSVENGCENACNAIPGAPRCQCPAGALQADRSCTA-SATQSCNDLCENFCVPNPDP 303
 DB 68 CAMDHNCEBOLCVNVPSCYSGYALADEGRKCAVAACASENHGCEHCY-NAD-- 124
 OY 304 GSYSCMCETGYRLADQHRCEVDVDDCTLEPSPCRPCVNTQGGFECHCYNYVL-VDSGC 362
 DB 125 GSYLCQCHGEFALNPDKTKTKIDYCASSNHGCONHECVNTDSDYSCHLKGFTLNPDKTK 184
 OY 363 VEPVPCF--RANCEYQCPINOTSTLYCAEGFAPIPH-----EPHRCQMF 408
 DB 185 CRRINCYALKPCGHEHCYNNME-SYCCRHRGYTLDPNCKTCSRVDHCAQODHGEQDC 243
 OY 409 NOTACPADCPNTOAS--CECPREYITLDDGF-ICTDIDEC--ENGFGCGVCHNLPGRF 462
 DB 244 -----LNTEDSFVCCQSSGEFLTNECLKTCSRVDYCLSLDHG--CEYSCVNDRSR 291
 OY 463 ECLGCPDPSALARIHTDCDSGKYDG--GDSC 491
 DB 292 ACQC-PGHHVLRSDGKTC--AKIDSCALSDHG 320

RESULT 13

A49457
 fibulin-2 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 11-Jan-2000

C:Accession: A49457; S74095

R:Pen, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.

J. Cell Biol. 123, 1269-1277, 1993

A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein with

A:Reference number: A49457; MIMD:94064787

A:Accession: A49457

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1221 <PAN>

A:Cross-references: GB:X75285; NID:9437046; PIDN:CAAS3040.1; PID:9437047

R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.

Eur. J. Biochem. 240, 427-434, 1996

A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix met

A:Reference number: S74094; MIMD:96439073

A:Accession: S74095

A:Molecule type: protein

A:Residues: 236-238, 'X', '240-247; 260-275; 336-344, 'L', '346-361; 405-426; 566-568, 'EW', '569-589

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer

F:942-978/Domain: EGF homology <EGF>

Query Match 12.2%; Score 357; DB 2; Length 1221;
 Best Local Similarity 29.5%; Pred. No. 4, 3e-14;
 Matches 123; Conservative 39; Mismatches 141; Indels 114; Gaps 27;

OY 154 EQQCEKAD-GEICEFFPATC-----RPLAV-----EPGAAAVS-----IT 192
 DB 521 EQQSCENPNLGYPCN-HVMSCEGEPPLIVPEVRRPEPEAPRRSEMEMASREALS 579
 OY 193 YET-----PFAARGAPFO--ALP-----VGSAAVAPGLDLM-----CTAPRG 229
 DB 580 LGTEAELEPNSLPBGDDQDECLMLPGELOHLCINTVGYTRCACPFGEFGDGRCPDRG 639
 OY 230 AVQGHMARE-APGAWDCSV-----ENGCEHACNAIPGAPRCQCPAGAA 272
 DB 640 APQLDTARESAFRRSASQVSNITPLVPQPNCKDNPCQVGRVVGDTAMCSCFFGYA 699
 OY 273 LDADRSC-----TASATQSCNDLCENFCVNPDPQPGSYSC-----MCEGYRLADQHR 322
 DB 700 IMADGVSCDEDECLMGTHDCS--WKQPCV--NTIGSFYCVNHTVLCAEGYILNA-HRK 753
 OY 323 CEDVDCTLEPSPC--PQRCVNTQGGFECH--CYRNYDLVDGCEVPVPCFRA--NC 374
 DB 754 CYDINECTYDLHTCTRAHCYNTQSGCYTALNCEPGYVLTDECTD-VDECYTGHNC 812
 OY 375 E--YQCPINOTSTLYCV--CAEGFAPIPH-----EPHRCQMFQNOTACPA 415
 DB 813 QAGFCQCN-TKGSFYQARQCMDFLODPGNGCVINDICTSLLEPCRSFGSCINTVGSY 871
 OY 416 DCDPNTQASCEPREGY-ITDGFICTDIDECENGFGCG--VGHNLPCTECICGR 468
 DB 872 TQQRNPLY--CGRGTANERBSCEVDVNECTGVHRCBSQLCTNLPGSTRCDCKP 925

RESULT 14

T09059
 notch4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jan-2000

C:Accession: T09059

R:Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Loretz, C.;
 submitted to the EMBL Data Library, October 1997

A:Description: Sequence of the mouse major histocompatibility locus class III region.

A:Reference number: Z16543

A:Accession: T09059

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1964 <ROW>

A:Cross-references: EMBL:AF030001; NID:92564945; PID:92564947

C:Genetics:

A:Gene: notch4

A:Map position: 17

A:Insertions: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1;
 1679/3; 1729/1; 1761/3

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: receptor; signal transduction

F:514-545/Domain: EGF homology <EGF>

Query Match 11.4%; Score 332.5; DB 2; Length 1964;
 Best Local Similarity 26.3%; Pred. No. 1, 8e-12;
 Matches 128; Conservative 34; Mismatches 142; Indels 183; Gaps 31;

OY 134 GFLCAVAASAEITVSEPRIMEEQCEVADDFICE-FHEPRTCPRLAVEGAAAANSTIR 192
 DB 35 GGTCTLRSLRGQIC-----QC--APGLDETQCPDPPCR----- 66
 OY 193 YGTPPAAGADFOAL--PVGSSAAVAPLGLQKCTAPPG--AVQGHMAREAPGAWDC 245
 DB 67 -DTQCKRNGSGQALNLPFRPSRSTSLTTPHPSCTCPSPSGTGRDQTHLELCPSPF-C 124

QY 246 SVENGGEGHACNAIPGAPRCQCPAGAAIQAIDGRSCTASATQSCNDLCEHFCVNPDPG- 304
 Db 125 S--NGG--HCYVQASGRPQSCSEPGMT-----GEQCQ-----LRDCSANPCAMGG 166
 QY 305 -----SYSCMEFYRLADQHRCE-DVDDCLLESPQRQ--RCVNTQSGFECHC--- 351
 Db 167 VCLATVPQIQRCPPGF---EGHTCEKRDINECFLEPGRPQGTSCNLTGSCJCLCPVG 222
 QY 352 --YPNYDLVDECEVPDPCEFRANCEYOCPL---NOTSYLCVCAEGFAPIPHE----- 400
 Db 223 QEGPQCKRKAKAC--PPGSCINLG--TCQLVPEGHSTPHLCCLPPGFTGLDEMNDDC 277
 QY 401 -PHRCOMFCNOTACPADCDPNTQASCECPREGIIDDGFICT-DIDCE-----NGGF 450
 Db 278 VRHOCQ---NGATCIDGLDITYT---CLCPKTV---KGMDCSEDIDECERAGPPRCRNGT 328
 QY 451 C-----SGVCHNLPGTFPCICGP----- 468
 Db 329 CONTAGSFHCVCVSGWAGAGEENLDDCAATCAPGSTCIDRVGSFSCCLPPGRTGLCH 388
 QY 469 --DSALAR--HIGTDCDSGKVDG-----GDSGS-----GEPPSP----- 499
 Db 389 LEDMCLSGPCHVNAQCSINPLTGSTLCIQPGYSGSTCHODLDECQMAQGPSPCEHGS 448
 QY 500 ---TPGS 503
 Db 449 CINTPGS 455

RESULT 15

T43210
 fibulin-1D precursor - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000

C:Accession: T43210

R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.

submitted to the EMBL Data Library, June 1998

A:Description: Identification of chicken and C. elegans fibulin-1 homologs and character

A:Reference number: Z22337

A:Accession: T43210

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-589 <BAR>

A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1

C:Genetics:

A:Note: Intron positions not resolved (incomplete sequence)

C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 11.4% Score 331; DB 2; Length 589;

Best local similarity 28.1% Pred. NO. 8.4e-13; Matches 126; Conservative 49; Mismatches 166; Indels 108; Gaps 28;

QY 103 GPLRGF-QWVTGDNNTSYSRMARLDLNGAPLGPVCYVSAEATVPSEPT-----WEEQ 156
 Db 19 GCLRSFNKCCMGDLEITH---ASELITGRPLNDPHVHLGDRCASSHCHLCHDRGGEKV 75
 QY 157 OCEVKA-----DGLFCEHFP--ATCRPLAVE-----PGAAAAVSTYGTFFA--- 198
 Db 76 ECSCRSRGLDAPDGMACYDHIDECATLMDCLIESQRLNTPGSKIRTLSCGTGAMDS 135
 QY 199 --ARGADFOALPYGSSAVALPLGIQLMCTAPPAVO-----GHMAREAPGAMDCS---- 246
 Db 136 ETERCRDYDECNLGSH---DCGLPYQCRNTQSGYRCDAAKKGSGDELONPMTGETSITC 191
 QY 247 -----VENGGCE-----HACNA-----IPGAPRCQ-----CPAGAAIQAIDGRSCTA 282
 Db 192 PNGYYPKNGMCDIDECVTGNHCGAGECVNTPGSPRCQCKGNLCANGYEYN----- 243
 QY 283 SATQSCNDL--CEH-----FCVNPDPQPGSYSCMCEYRLADQHRCEVDYDCIL--- 331
 Db 244 GATGCEDEVNECQOGVCGSMCEI---NLPGTYKCKGPGYEFENDAKKRCEDEVDECIKFAG 300

QY 332 EPSPCPQRCVNTQGGFECHCYPNYDLV--DGECEVPDPCER--ANCEYOCPLNQTSYLC 388
 Db 301 HVCDSLAEICINTIGSFCECKKRGFQLASDGRCEDEVNCTTGIAECQKCVNI--PGSYQC 359
 QY 389 VCAEGFAPIP-----HEPHRCOMEC--NOTACPADCDPNTQAS--CCCPGSGY--IIDDGFI 438
 Db 360 ICDRGFALGPDTGCKEDIDECISIMAGSNDICMGCC--INTKGYLCCPPGRTKIQPDGRT 418
 QY 439 CTYDIDCEENGFCGSG---VCHNLPGTFPC 464
 Db 419 CVYDVECAMGE--CAQSDKVCVNTIGSFKC 446

Search completed: May 8, 2002, 12:40:05
 Job time: 251 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:38:52 ; Search time 42.86 Seconds

(without alignments)
891.783 Million cell updates/sec

Title: US-09-509-994-2
Perfect score: 2916
Sequence: 1 MGVIVLGLALALGLGFAP.....PSPTPGSTTLTPAVGLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-----------------------------|
| 1 | 2916 | 100.0 | 516 | 20 | AAV09348 Human thrombomodul |
| 2 | 2916 | 100.0 | 516 | 21 | AAV83935 Human thrombomodul |
| 3 | 2912 | 99.9 | 516 | 13 | AAK22016 Truncated human th |
| 4 | 2912 | 99.9 | 516 | 20 | AAV09347 Human thrombomodul |
| 5 | 2912 | 99.9 | 516 | 21 | AAV83934 Human thrombomodul |
| 6 | 2912 | 99.9 | 516 | 21 | AAV69529 Human thrombomodul |
| 7 | 2912 | 99.9 | 516 | 12 | AAK11534 Human thrombomodul |
| 8 | 2912 | 99.9 | 516 | 14 | AAK41806 Human thrombomodul |
| 9 | 2912 | 99.9 | 516 | 14 | AAK43031 Human thrombomodul |
| 10 | 2910 | 99.8 | 515 | 12 | AAK10617 Soluble thrombomod |
| 11 | 2910 | 99.8 | 575 | 9 | AAK2070 Human thrombomodul |

| | | | | | |
|----|--------|------|-----|----|-----------------------------|
| 12 | 2908 | 99.7 | 516 | 13 | AAK22018 Human thrombomodul |
| 13 | 2908 | 99.7 | 516 | 13 | AAK31572 Human thrombomodul |
| 14 | 2904 | 99.6 | 516 | 14 | AAK22017 Human thrombomodul |
| 15 | 2903 | 99.6 | 516 | 13 | AAK20639 Human urinary thro |
| 16 | 2879 | 98.7 | 575 | 20 | AAW73970 Human thrombomodul |
| 17 | 2862 | 98.1 | 575 | 13 | AAK22189 Sequence of thromb |
| 18 | 2826 | 96.9 | 498 | 18 | AAK84185 Human derived thro |
| 19 | 2826 | 96.9 | 498 | 18 | AAW01600 Human thrombomodul |
| 20 | 2826 | 96.9 | 498 | 21 | AAV67402 Novel sugar chain- |
| 21 | 2824 | 96.8 | 497 | 17 | AAK94607 Human recombinant |
| 22 | 2820 | 96.7 | 497 | 19 | AAW69520 rSTM protein SEQ I |
| 23 | 2807 | 96.3 | 500 | 21 | AAV69530 Novel sugar chain- |
| 24 | 2804 | 96.2 | 494 | 21 | AAV67401 Novel sugar chain- |
| 25 | 2766 | 94.9 | 494 | 16 | AAK78726 Recombinant thromb |
| 26 | 2764 | 94.8 | 494 | 13 | AAK24400 Thrombin-binding s |
| 27 | 2749 | 94.3 | 494 | 16 | AAK78727 Thrombin-binding s |
| 28 | 2744 | 94.1 | 480 | 13 | AAK18877 Truncated human th |
| 29 | 2716 | 93.1 | 480 | 13 | AAK22013 Mature thrombomod |
| 30 | 2686 | 92.1 | 475 | 13 | AAK22032 Modified thrombomo |
| 31 | 2680 | 91.9 | 476 | 16 | AAK78725 Modified thrombomo |
| 32 | 2680 | 91.9 | 476 | 16 | AAK86376 Sequence of human |
| 33 | 2678 | 91.8 | 476 | 16 | AAK86377 Thrombomodulin an |
| 34 | 2661 | 91.3 | 535 | 9 | AAK80641 Thrombomodulin an |
| 35 | 2623 | 90.0 | 462 | 15 | AAK45336 Thrombomodulin an |
| 36 | 2619 | 89.8 | 462 | 15 | AAK45337 Thrombomodulin an |
| 37 | 2619 | 89.8 | 462 | 15 | AAK45338 Thrombomodulin an |
| 38 | 2602 | 89.2 | 462 | 15 | AAK45339 Thrombomodulin an |
| 39 | 2600 | 89.2 | 462 | 15 | AAK45350 Thrombomodulin an |
| 40 | 2595 | 89.0 | 462 | 15 | AAK45343 Thrombomodulin an |
| 41 | 2576 | 88.3 | 462 | 15 | AAK45349 Thrombomodulin an |
| 42 | 2575 | 88.3 | 462 | 15 | AAK45353 Thrombomodulin an |
| 43 | 2570 | 88.1 | 462 | 15 | AAK45347 Thrombomodulin an |
| 44 | 2566.5 | 88.0 | 461 | 15 | AAK45335 Thrombomodulin an |
| 45 | 2565 | 88.0 | 462 | 15 | AAK45346 Thrombomodulin an |

ALIGNMENTS

RESULT 1
ID AAY09348 standard; Protein; 516 AA.
AC AAY09348:
XX 08-JUL-1999 (first entry)
XX Human thrombomodulin SEQ ID NO:2.
XX DE Human: thrombomodulin; aqueous parenteral solution; storage;
KW distribution; acute coronary syndrome; thrombosis; embolism;
XX diabetes.
XX OS Homo sapiens.
XX XX W09918994-A1.
XX XX 22-APR-1999.
XX PF 13-OCT-1998; 98WO-JP04609.
XX PR 11-NOV-1997; 97JP-0308523.
XX PR 15-OCT-1997; 97JP-0281659.
XX PA (ASAH) ASAH KASEI KOGYO KK.
XX PI Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M,
XX WPI: 1999-277444/23.
XX DR N-PSDB; AAX55880.
XX Stable aqueous parenteral thrombomodulin solution - comprising
PT buffer and surfactant, useful for treating acute coronary syndrome,

PT thrombosis, embolism, and diabetes
XX
PS Claim 7; Page 90-92; 97pp; Japanese.
XX

CC The present invention describes a method for maintaining the quality of
CC an aqueous parenteral solution of thrombomodulin comprising buffer and
CC surfactant aseptically filled in a case or syringe. Maintaining the
CC quality of an aqueous, parenteral thrombomodulin solution is
CC characterised in that the solution: (a) comprises soluble thrombomodulin,
CC a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically
CC filled into: (i) a case or (ii) a syringe without any empty space; and
CC (c) is kept in liquid form in storage and distribution and not frozen or
CC freeze dried. Thrombomodulin is used to treat acute coronary syndrome
CC (such as myocardial infarction, unstable angina and coronary artery
CC blockage), thrombosis (e.g. cerebral, vascular and peripheral blood
CC vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.
CC Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic
CC lupus erythematosus or Barrett's syndrome), multiple organ failure,
CC disseminated intravascular coagulation, transient ischaemia, diabetes,
CC liver veno-occlusive diseases and deep vein thrombosis. The composition
CC is stable for a long period of time and can be stored and distributed in
CC ready to use form avoiding the problems of dissolution and accuracy when
CC preparing on demand. The present sequence represents human
CC thrombomodulin.
XX
SQ

Sequence 516 AA:

Query Match 100.0%; Score 2916; DB 20; Length 516;
Best Local Similarity 100.0%; Pred. No. 4.4e-163;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLVYLGLALAGLGFPAPEPQPGSGQVEHDCFALYPPATFLNASQICDGLRGHLM 60
DB 1 mlgvlylgatalaglgfpapeepqpgsgqvehdcfalyppatflnaasqicdglrghlm 60
QY 61 TVRSSVAADYISLLNGDGVGRRRLMIGQLPPGCGDPRRLGLPGFQVWTGDNNTSYS 120
DB 61 tvrssvaadvysllngdgvgrrrlmgqlppgcgdprkrlglpgfqvwtgdnntsys 120
QY 121 RMARLDLNGAPLCGPLCVANSAAEATVPSEPIWEBOQCEVKADGFLCEHFPAATCRPLAV 180
DB 121 rmarldlngaplcgplcvansaeeatvpsepiweeqcevkadgflcehfpatcrplav 180
QY 181 EPGAAAAVSYITGTPFAARGADFOALPVGSSAAVAPLGIQLMCTAPPGAVGCHMARBP 240
DB 181 epgaaaavsyitgtpfaargadfoalpvssaaavaplgqlmctappgavqghmarbp 240
QY 241 GAWDCSVENGCEHACNAIPGARPCCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcehacnaipgarpcqpagaaqadgrsctasaatqscndlcehfcvnp 300
QY 301 DPGSTSCMCETGYRLAADQHRCEVDVDCILBSPCPQRCVNTQSGFECHCYPNYDLVNG 360
DB 301 dpgstyscmcetgyrllaadqhrcevdvdcilbepsqpcvntqsgftechcypnydlvng 360
QY 361 ECEVPDPCFRANCFEYQCPPLNOTSTLYCABGFAPRIHPEHRCQMFQNOTACPADCDPN 420
DB 361 ecevpdpfrancfeyqcpplnotstlycabgfaprihpehrcqmfqnotacpadcdpn 420
QY 421 TQASCECPGYIIDDGFICTDIDECENGFGSGVCHNLGTFCITCGPSALARHIGTDC 480
DB 421 tqascecpgyiiddgfictdidecengfgsgvchnlpgtfecltqpsalarhigtgc 480
QY 481 DSGKYDGDGSGGSEPPSPTPGSTLTPPAVGLVHSG 516
DB 481 dsgkydgdsgsgseppspptpgstltpavglvhs 516

RESULT 2
ID AAY83935
XX AAY83935 standard: Protein; 516 AA.

AC AAY83935;
XX
XX 28-JUL-2000 (first entry)
DE Human thrombomodulin TMD protein #2.
XX
XX Human; thrombomodulin; vasculitis; protein C; thrombin.
XX
XX Homo sapiens.
XX
XX JP2000053582-A.
XX
XX 22-FEB-2000.
XX
XX 06-AUG-1998; 98UP-0222688.
XX
XX 06-AUG-1998; 98JP-0222688.
XX
XX (ASAH) ASAH KASEI KOGYO KK.
XX
XX WPI; 2000-353249/31.
XX
XX N-PSDB; AAA10028.
XX

PT Treating agent for vasculitis contains peptide which promotes
PT activation of protein C by thrombin -
PS Claim 4; Page 13-14; 18pp; Japanese.
XX
XX This sequence represents a human thrombomodulin protein. The invention
XX relates to a treating agent for vasculitis containing a peptide which
XX promotes activation of protein C by thrombin.
XX

Sequence 516 AA:

Query Match 100.0%; Score 2916; DB 21; Length 516;
Best Local Similarity 100.0%; Pred. No. 4.4e-163;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLVYLGLALAGLGFPAPEPQPGSGQVEHDCFALYPPATFLNASQICDGLRGHLM 60
DB 1 mlgvlylgatalaglgfpapeepqpgsgqvehdcfalyppatflnaasqicdglrghlm 60
QY 61 TVRSSVAADYISLLNGDGVGRRRLMIGQLPPGCGDPRRLGLPGFQVWTGDNNTSYS 120
DB 61 tvrssvaadvysllngdgvgrrrlmgqlppgcgdprkrlglpgfqvwtgdnntsys 120
QY 121 RMARLDLNGAPLCGPLCVANSAAEATVPSEPIWEBOQCEVKADGFLCEHFPAATCRPLAV 180
DB 121 rmarldlngaplcgplcvansaeeatvpsepiweeqcevkadgflcehfpatcrplav 180
QY 181 EPGAAAAVSYITGTPFAARGADFOALPVGSSAAVAPLGIQLMCTAPPGAVGCHMARBP 240
DB 181 epgaaaavsyitgtpfaargadfoalpvssaaavaplgqlmctappgavqghmarbp 240
QY 241 GAWDCSVENGCEHACNAIPGARPCCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcehacnaipgarpcqpagaaqadgrsctasaatqscndlcehfcvnp 300
QY 301 DPGSTSCMCETGYRLAADQHRCEVDVDCILBSPCPQRCVNTQSGFECHCYPNYDLVNG 360
DB 301 dpgstyscmcetgyrllaadqhrcevdvdcilbepsqpcvntqsgftechcypnydlvng 360
QY 361 ECEVPDPCFRANCFEYQCPPLNOTSTLYCABGFAPRIHPEHRCQMFQNOTACPADCDPN 420
DB 361 ecevpdpfrancfeyqcpplnotstlycabgfaprihpehrcqmfqnotacpadcdpn 420
QY 421 TQASCECPGYIIDDGFICTDIDECENGFGSGVCHNLGTFCITCGPSALARHIGTDC 480
DB 421 tqascecpgyiiddgfictdidecengfgsgvchnlpgtfecltqpsalarhigtgc 480
QY 481 DSGKYDGDGSGGSEPPSPTPGSTLTPPAVGLVHSG 516
DB 481 dsgkydgdsgsgseppspptpgstltpavglvhs 516

Db 481 dsqkvdgdsqsgpppsptltpavglvhsq 516

RESULT 3

AA022016

ID AAR22016 standard; Protein; 516 AA.

XX

AC AAR22016;

XX

DT 03-JUL-1992 (first entry)

XX

DE Truncated human thrombomodulin encoded by plasmid pSV2TMD1.

XX

KW Thrombin binding site; blood clotting; TMD1 deleter.

XX

OS Homo sapiens.

XX

PN EP474273-A.

XX

PD 11-MAR-1992.

XX

PE 05-AUG-1991; 91EP-0202009.

XX

PR 03-AUG-1990; 90JP-0204978.

XX

PA (ASAH) ASAH KASEI KOGYO.

XX

PI Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;

XX

DR WPI; 1992-081820/11.

XX

PT New polypeptide inhibits blood coagulation and platelet

PT aggregation - promotes thrombin catalysed protein C activation

PT for treating myocardial infarction, thrombosis, embolism, etc.

XX

PS Example 1; Page 17; 112pp; English.

XX

CC Plasmid M13mp19TMD1 (see AAR22013) was used as template for

CC site-directed mutagenesis. A 177bp fragment was deleted using the

CC "TMD1 deleter" oligonucleotide to give plasmid M13TMD1 which encodes

CC the first 516 N-terminal amino acids of human thrombomodulin. Plasmid

CC M13TMD1 was completely digested with HindIII and BamHI and a TMD1

CC fragment of ca. 1700bp was isolated. The fragment was ligated to

CC HindIII- and BglII-cut plasmid pSV2-dhfr to give the recombinant

CC plasmid pSV2TMD1. See AAR22014-R22022 and AA025072.

XX

SO Sequence 516 AA;

Query Match 99.9%; Score 2912; DB 13; Length 516;
Best Local Similarity 99.8%; Pred. No. 7.5e-163;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MGVLYLGAIALAGLPAPAPQPGSGQCEVHDCFLALYCPATFELNASOICDGLRGLHM 60
DB 1 mlyglvlgalaiaeglfgpapapqpgsgqcevhdcfalygpafellnasqicdglrghlm 60

OY 61 TVRBSVADYVSLINDGCVGRRLMIGQLPRGCGDPRRLGRLRFQWWTGNNNTSYS 120
DB 61 tvrsvadvysllndgvgvrrrlvlgqlpprgcgdprrlgqlrfqwtggnntsys 120

OY 121 RWARLDNAGAPLCPGPLCAVAASAEATVPSEPIWEQCEVAKDCLFEHFPATCRPLAV 180
DB 121 rwarldnagaplcpgplcvaasaeatvpsepiweqcevakdclfehfpatcrplav 180

OY 181 EPGAAAASVTYGTTPPAACADFOALPVGSSAAVADLGLQMLCTAPGAVQGHAREAP 240
DB 181 epgaaaasvtygttppaacadfoalpvgssaaavadlglqlmctapgavqghwareap 240

OY 241 GAWCSEVNGGCEHACNARIGAPRCQCPAGALQADGRSCTASTQSCNLCHEHFCVPM 300
DB 241 gawcsevnggcehacnarpaprcqcpagalaqadgrsctastqscnldchehfcvpm 300

OY 301 DQPGSYSCMCEFGYRLAADOHRCEDVDICILEPSPQRCVNTQGFECCHCYPNYLDYDC 360
DB 301 dqpgsyscmcefgyrlaadohrcedvddcillepspqrcvntqgfecchcypnyldvdc 360

OY 361 ECVPEYDPCFRANCEYQCQPLNQTSLYCAAGFAPIRPHPRRCQMFQNTACPADCCPN 420
DB 361 ecvpeydpfranceyqcqplnqtslycvaagfapirphprcqmfncqtacpadccpn 420

OY 421 TQASCCEPEGYIIDDGFCITDIDECENGFGSCGVCHNLGTFECTCGPDSALARIHGTDC 480
DB 421 tqasccepegyiiddgfcitdiddecengfgscgvchnlgtfectcgpdsalvrihgtdc 480

OY 481 DSGKVDGSDSGSGEPPSPPTPSSTLTPPAVGLVHSG 516
DB 481 dsqkvdgdsqsgpppsptltpavglvhsq 516

RESULT 4

AA09347

ID AAY09347 standard; Protein; 516 AA.

XX

AC AAY09347;

XX

DT 08-JUL-1999 (first entry)

XX

DE Human thrombomodulin SEQ ID NO:1.

XX

KW Human; thrombomodulin; aqueous parenteral solution; storage;

KW distribution; acute coronary syndrome; thrombosis; embolism;

KW diabetes.

XX

OS Homo sapiens.

XX

PN WO918994-A1.

XX

PD 22-APR-1999.

XX

PF 13-OCT-1998; 98WO-JP04609.

XX

PR 11-NOV-1997; 97JP-0308523.

XX

PR 15-OCT-1997; 97JP-0281659.

XX

PA (ASAH) ASAH KASEI KOGYO KK.

XX

PI Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;

XX

DR WPI; 1999-277444/23.

DR N-PSDB; AAX5879.

XX

Stable aqueous parenteral thrombomodulin solution - comprising
PT buffer and surfactant, useful for treating acute coronary syndrome,
PT thrombosis, embolism, and diabetes

XX

PS Claim 6; Page 87-89; 97pp; Japanese.

XX

The present invention describes a method for maintaining the quality of
CC an aqueous parenteral solution of thrombomodulin comprising buffer and
CC surfactant aseptically filled in a case or syringe. Maintaining the
CC quality of an aqueous, parenteral thrombomodulin solution is
CC characterised in that the solution: (a) comprises soluble thrombomodulin,
CC a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically
CC filled into: (i) a case or (ii) a syringe without any empty space; and
CC (c) is kept in liquid form in storage and distribution and not frozen or
CC freeze dried. Thrombomodulin is used to treat acute coronary syndrome
CC (such as myocardial infarction, unstable angina and coronary artery
CC blockage), thrombosis (e.g. cerebral, vascular and peripheral blood
CC vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.
CC Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic
CC lupus erythematosus or Bartlett's syndrome), multiple organ failure,
CC disseminated intravascular coagulation, transient ischaemia, diabetes,
CC liver veno-occlusive diseases and deep vein thrombosis. The composition
CC is stable for a long period of time and can be stored and distributed in
CC ready to use form avoiding the problems of dissolution and accuracy when

CC preparing on demand. The present sequence represents human
 CC thrombomodulin.
 XX
 SQ Sequence 516 AA.

Query Match 99.9%; Score 2912; DB 20; Length 516;
 Best Local Similarity 99.8%; Pred. No. 7.5e-163;
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLYLGALALAGLPAPAEQPGSGQVCHDCEALYPGATFLNLSQICDGLRHLM 60
 |||||||
 DB 1 mglvlylgalaalaglpapaeqpqsgvchdcfalyppatflnlsqicdglrhlm 60
 QY TVRSSVADYISLLNGDGVGRRLMIGLOLPFGCGDKRLGPIRGFQWVGDNNTSYS 120
 |||||||
 DB 61 tvrssvaadvyslllmgdgvgrrrlmglolpfgcgdkrlgpirgfwvgdnntsys 120
 QY 121 RMARLDLNGAPLGGPLCVASAATVPSEPIWEQCEVKADGFLCEHFPAICRPLAV 180
 |||||||
 DB 121 rmarldlngaplgplcvasaaatvpsepiweeqcevkadgflcehfpaicrplav 180
 QY 181 EPGAANAASITGTGPPAARGADFOALPVSSAANAAPLGIOLMCTAPGAVOGHMAAREAP 240
 |||||||
 DB 181 epgaanaasitgtgppaargadfoalpvssanaaaplgilolmctappgavghmareap 240
 QY 241 GAMDCEVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
 |||||||
 DB 241 gamdcevnggcehacnaipgaprcqcpagalaqadgrsctasatqscndlcehfcvnp 300
 QY 301 DQSGSYSCMCEGTGRLAADHRCEDVDCLERSPCQRCVNTGGFEGHCYPRNTDLYDG 360
 |||||||
 DB 301 dqsgyscmcegtgrylaadhrcedvdclleerspqrctvntggfeghcyprnydlvdg 360
 QY 361 ECEVPDPCFRANCEYOCPLNOTSYLCVCAEGFAPLPHBPHRCQMFQNOTACPADCDPN 420
 |||||||
 DB 361 ecevpdpccfranceyqcpnlngtsylcvcaegfaplhpbphrcqmfqnotacpadcdpn 420
 QY 421 TQASCECEPEGYIIDDGFICTDIDECENGFGSGVCHNLPGTEECICGPDALAHITGDC 480
 |||||||
 DB 421 tqascecepegyiiddgfiectdidceengfgsgvchnlpgteecicgpdalahrhtgdc 480
 QY 481 DSGKVDGDSGSGRPPSPPTPGSTLTPPAVGLVHSG 516
 |||||||
 DB 481 dsgrvddgdsgrppspptpgstltppavglvhsq 516

RESULT 5
 AAY83934
 ID AAY83934 standard; Protein: 516 AA.
 XX
 AC AAY83934;
 XX
 DT 28-JUL-2000 (first entry)
 XX
 DE Human thrombomodulin TMD protein.
 XX
 KW Human; thrombomodulin; vasculitis; protein C; thrombin.
 XX
 OS Homo sapiens.
 XX
 PN JP2000053582-A.
 XX
 PD 22-FEB-2000.
 XX
 PF 06-AUG-1998; 98JP-0222688.
 XX
 PR 06-AUG-1998; 98JP-0222688.
 XX
 PA (ASAH) ASAH KASEI KOGYO KK.
 XX
 WI WI; 2000-353249/31.
 DR N-PSDB; AAA10027.

XX
 PT Treating agent for vasculitis contains peptide which promotes
 XX activation of protein C by thrombin -
 XX
 PS Claim 4; Page 10-12; 18pp; Japanese.

CC This sequence represents a human thrombomodulin protein. The invention
 CC relates to a treating agent for vasculitis containing a peptide which
 CC promotes activation of protein C by thrombin.
 XX

SQ Sequence 516 AA:

Query Match 99.9%; Score 2912; DB 21; Length 516;
 Best Local Similarity 99.8%; Pred. No. 7.5e-163;
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLYLGALALAGLPAPAEQPGSGQVCHDCEALYPGATFLNLSQICDGLRHLM 60
 |||||||
 DB 1 mglvlylgalaalaglpapaeqpqsgvchdcfalyppatflnlsqicdglrhlm 60
 QY 61 TVRSSVADYISLLNGDGVGRRLMIGLOLPFGCGDKRLGPIRGFQWVGDNNTSYS 120
 |||||||
 DB 61 tvrssvaadvyslllmgdgvgrrrlmglolpfgcgdkrlgpirgfwvgdnntsys 120
 QY 121 RMARLDLNGAPLGGPLCVASAATVPSEPIWEQCEVKADGFLCEHFPAICRPLAV 180
 |||||||
 DB 121 rmarldlngaplgplcvasaaatvpsepiweeqcevkadgflcehfpaicrplav 180
 QY 181 EPGAANAASITGTGPPAARGADFOALPVSSAANAAPLGIOLMCTAPGAVOGHMAAREAP 240
 |||||||
 DB 181 epgaanaasitgtgppaargadfoalpvssanaaaplgilolmctappgavghmareap 240
 QY 241 GAMDCEVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
 |||||||
 DB 241 gamdcevnggcehacnaipgaprcqcpagalaqadgrsctasatqscndlcehfcvnp 300
 QY 301 DQSGSYSCMCEGTGRLAADHRCEDVDCLERSPCQRCVNTGGFEGHCYPRNTDLYDG 360
 |||||||
 DB 301 dqsgyscmcegtgrylaadhrcedvdclleerspqrctvntggfeghcyprnydlvdg 360
 QY 361 ECEVPDPCFRANCEYOCPLNOTSYLCVCAEGFAPLPHBPHRCQMFQNOTACPADCDPN 420
 |||||||
 DB 361 ecevpdpccfranceyqcpnlngtsylcvcaegfaplhpbphrcqmfqnotacpadcdpn 420
 QY 421 TQASCECEPEGYIIDDGFICTDIDECENGFGSGVCHNLPGTEECICGPDALAHITGDC 480
 |||||||
 DB 421 tqascecepegyiiddgfiectdidceengfgsgvchnlpgteecicgpdalahrhtgdc 480
 QY 481 DSGKVDGDSGSGRPPSPPTPGSTLTPPAVGLVHSG 516
 |||||||
 DB 481 dsgrvddgdsgrppspptpgstltppavglvhsq 516

RESULT 6
 AAY69529
 ID AAY69529 standard; Protein: 516 AA.
 XX
 AC AAY69529;
 XX
 DT 10-APR-2000 (first entry)
 XX
 DE Human thrombomodulin variant, SEQ ID NO:3.
 XX
 KW Thrombomodulin; TM; soluble; affinity purification; cation exchange;
 XX anticoagulant; thrombolytic.
 XX
 OS Homo sapiens.
 XX
 PN JP113411990-A.
 XX
 PD 14-DEC-1999.

PF 23-MAR-1999; 99JP-0077518.
XX
PR 30-MAR-1998; 98JP-0084389.
XX
PA (ASAHI) ASAHI KASEI KOGYO KK.
XX
DR WPI: 2000-101696/09.
DR N-PSDB; AA255965.
XX
PT Preparation of highly pure soluble thrombomodulin - used as an
PT antiblood coagulation agent and a thrombolytic agent
XX
XX
PS Claim 9; Page 30-32; 36pp; Japanese.
XX
XX
CC The invention relates to a novel method for the preparation of highly
CC pure soluble thrombomodulin (TM) containing substantially no serum-
CC derived or antibody-derived substance. The method comprises isolating
CC soluble TM from unpurified serum via affinity purification using an
CC anti-TM antibody. The soluble TM is then treated with a cation
CC exchanger at a specific conductivity of 25 to 34 ms/cm and a pH of 3 to
CC 4, and the fraction containing the soluble TM is isolated. The soluble
CC TM can be used as a blood anticoagulation agent and a thrombolytic agent.
CC This sequence represents a human thrombomodulin variant, designated
XX SEQ ID NO:3.
XX
SO Sequence 516 AA;

QY 301 DPGSYSCMCEGTGYRLADQHRCEVDVDCILPSPCPCRCVNTQGGFEGHCYPNYDLVDS 360
 Db 301 dpgsysscmetcgyrlaadqhrcevdvdcilpsspccrcvntqggfeghcyppydlvds 360
 QY 361 ECVPEVDPGFRANCERYCQPLNNTSYLCVCAEGFAPLPHPHRCQMFQNTACRPADCPN 420
 Db 361 ecvpevdpfranceygcqplnntsylycvcaegfaplhphrcmqmfncqtaacpdcn 420
 QY 421 TQASCEPPEGYLLDDGFICTDIDECENGFCGVCNHLPGTFECICGSPSALAHIGTDC 480
 Db 421 tqascepegyllddgfictdidecengfcsgvchnlpgtfecicgspalsalvrlhgtdc 480
 QY 481 DSGKVDGSDSGSEPPSPPTGSLTPRPANGLVHSG 516
 Db 481 dsgkvdgdsqsgseppspptgsltlppavglvhs 516

RESULT 8
 AAR41806
 ID AAR41806 standard; peptide; 575 AA.
 AC AAR41806;
 XX 30-MAR-1994 (first entry)
 DT Thrombomodulin.
 DE Thrombomodulin.
 XX Transformation; fungus: blood coagulation; prevention; platelet;
 KM aggregation; thrombolytic activity; thrombolysis; human thrombomodulin;
 KM site-directed mutagenesis; promotion; protein C; activation; thrombin.
 XX Homo sapiens.
 OS JP05213998-A.
 PN 24-AUG-1993.
 PD 02-AUG-1991; 91JP-0282369.
 PE 03-AUG-1990; 90JP-0204978.
 PR 30-JUL-1991; 91JP-0189984.
 XX (ASAHI) ASAHI CHEM IND CO LTD.
 PA WPI: 1993-299652/38.
 DR Novel polypeptide obtd. by culturing transformed fungus - having
 PT blood coagulation preventing, platelet aggregation preventing and
 PT thrombolytic activities
 XX Disclosure; Fig 55; 65pp; Japanese.
 PS Novel polypeptides, obtd. by culturing transformed fungus, have
 CC blood coagulation preventing, platelet aggregation preventing
 CC and thrombolytic activities.
 CC In an example, plasmid M13mp19TMD3 (constructed from pSV2TMD2
 CC (ATCC 67238) contg. a human thrombomodulin sequence) was subjected
 CC to site directed mutagenesis to prepare pSV2TMD7. Plasmid pSV2TMD7
 CC was transfected to COS-1 cells. The activity of promoting protein C
 CC activation by thrombin of the peptide produced by the transformed
 CC COS-1 cell was measured. The amt. of the peptide was determined.
 XX Sequence 575 AA;

Query Match 99.98; Score 2912; DB 14; Length 575;
 Best Local Similarity 99.88; Pred. No. 8.3e-163;
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVVIVGALALAGLPAPAPAPQPGSGSCVEHDCFALYGPATFLNASQICDGLRGLM 60
 Db 1 mlgviviagalalaglpapapapqpgsgscvehdcfalygpattfnasqicdglrghm 60

QY 61 TVRSSVAADVLSILNCDGGVRRRLWLGLOLPFGCGDPRFLRGLPQWMTGDNNTSYS 120
 Db 61 tvrssvaadvlsilnngdyvrrrlwlglolpfpgcgdprrlrglpqfwmtgdnntsys 120
 QY 121 RPARLDLNCAPLCGFLCVAVSAEAATVPSEPTWEEOCEKADGFLCEHFPATCRPLAV 180
 Db 121 rparldlncaplcgflcvavsaeeaavtvpseptweeocekadgflcehfpatcrplav 180
 QY 181 EPGAAAAAIVTGYTPFARAGADFOALPVGSSAAVAFGLQMLCTAPPGAVGSHMARAP 240
 Db 181 epgaaaaaivtytpfaragadfoalpvvgssaaavafglqmlctappgavghmarap 240
 QY 241 GAMDCSVENGCGEHCNMLIPGAPRCQCPAGALADQDRCTASTATQSCNDLCEHCVNRP 300
 Db 241 gamdcsvengcgehcncmlipgaprcqcpagalaadqdrctastatqscndlcehcvnrvp 300
 QY 301 DPGSYSCMCEGTGYRLADQHRCEVDVDCILPSPCPCRCVNTQGGFEGHCYPNYDLVDS 360
 Db 301 dpgsysscmetcgyrlaadqhrcevdvdcilpsspccrcvntqggfeghcyppydlvds 360
 QY 361 ECVPEVDPGFRANCERYCQPLNNTSYLCVCAEGFAPLPHPHRCQMFQNTACRPADCPN 420
 Db 361 ecvpevdpfranceygcqplnntsylycvcaegfaplhphrcmqmfncqtaacpdcn 420
 QY 421 TQASCEPPEGYLLDDGFICTDIDECENGFCGVCNHLPGTFECICGSPSALAHIGTDC 480
 Db 421 tqascepegyllddgfictdidecengfcsgvchnlpgtfecicgspalsalvrlhgtdc 480
 QY 481 DSGKVDGSDSGSEPPSPPTGSLTPRPANGLVHSG 516
 Db 481 dsgkvdgdsqsgseppspptgsltlppavglvhs 516

RESULT 9
 AAR43031
 ID AAR43031 standard; protein; 575 AA.
 AC AAR43031;
 XX 16-MAY-1994 (first entry)
 DT Human thrombomodulin.
 DE Human thrombomodulin.
 XX Anticoagulant; platelet aggregation inhibitor;
 KM protein C; activation; thrombin; thrombomodulin;
 KM coagulation disorder; thrombosis; myocardial infarction;
 KM embolism; telangiectasis; arteriosclerosis obliterans;
 KM disseminated intravascular coagulation; DIC; angina pectoris;
 KM gestosis; transient ischaemic attack.
 XX Homo sapiens.
 OS WO9322447-A.
 PN 11-NOV-1993.
 PD 30-APR-1993; 93WO-JP00578.
 PE 01-MAY-1992; 92JP-0112903.
 PR (ASAHI) ASAHI CHEM IND CO LTD.
 PA (ASAHI) ASAHI KASEI KOGYO KK.
 XX Kondo S, Toma K, Zushi M;
 DR WPI: 1993-368806/46.
 PT Peptide with anticoagulant and platelet aggregation inhibitor
 PT activity - which promotes protein C activation by thrombin and is
 PT useful in treating coagulation disorders e.g. thrombosis
 XX Disclosure; Fig 1; 84pp; Japanese.

XX New peptides (see AAR50069) are inhibitors of the blood coagulation
CC and platelet aggregation activities of thrombin and promote the
CC protein-C activation effect of thrombin. They can be produced
CC efficiently in pure form by culture of appropriate transformants,
CC and are useful in treatment of circulatory disorders such as
CC myocardial infarction, thrombosis, embolism, telangiectasis,
CC arteriosclerosis obliterans, disseminated intravascular
CC coagulation, angina pectoris, gestosis and transient ischaemic
CC attack.

Db 361 ecvepvdpccfranceygcqplnqtsylcvcaeafiapihpncqmfncqtaacpdcqpn 420
 QY 421 TQASCCEPEGYILDGDFICTDIDECENGCGSCVCHNLDEFTFCICGPPDSALARIHGTDC 480
 Db 421 tqascecepegyilddgfictdidecengcgscvchnlpgtfecicgppdsalarhgtc 480
 QY 481 DSGKVDGSDSGSEPPSPPTGSLTPPAVGLVHS 515
 Db 481 dsgkvdgsgdsgeppspptgsltpavglvhs 515

RESULT 11

AAp82070
 ID AAP82070 standard; protein; 575 AA.
 AC AAP82070;
 XX 19-OCT-1990 (first entry)
 DE Human thrombomodulin encoded by plasmid p2.1.
 KW thrombomodulin activity; protein C; anticoagulant;
 KW epidermal growth factor (EGF) domains.
 OS synthetic.
 XX
 PN W08809811-A.
 XX
 PD 15-DEC-1988.
 XX
 PE 09-JUN-1988; 88WO-DK00089.
 XX
 PR 12-JUN-1987; 87DK-0002990.
 XX
 PA (NOVO) NOVO INDUSTRI A/S.
 XX
 PI Nexo BA, Esper B;
 XX
 DR WPI: 1988-368626/51.
 DR N-PSDB; AAP82026.
 XX
 PT Recombinant protein having thrombomodulin activity -
 PT used in the therapeutic control of coagulation and the treatment
 PT and prevent of thrombotic episodes
 PS Disclosure; ; P; English.
 XX
 CC Protein is encoded by plasmid p2.1 derived from human cell
 CC line A549 known to express about 10000 molecules of thrombomodulin
 CC per cell. Plasmid p2.1 showed a strong hybridisation signal with a
 CC 60-mer bovine thrombomodulin probe. Thrombomodulins are used to
 CC potentiate a patient's anticoagulant capacity.
 CC See also AAN82027 and AAN82037.
 CC
 CC
 CC
 SQ Sequence 575 AA;

Query Match 99.8%; Score 2910; DB 9; Length 575;
 Best Local Similarity 99.8%; Pred. No. 1,1e-162;
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVLVIGALAGLGFAPAEPPQSGSCVEHDCFALYPCPATFLNASQICDGLRGHLM 60
 Db 1 mlgvlvlgalaaglgfapaeppqsgscvehdcfalyppatflnasqicdglrghlm 60
 QY 61 TVRSSVAADVISLLNSDGVGRRRLWIGLQLEPGCGDKRLGRLGFGQWVGDNNTSYS 120
 Db 61 tvrssvaadvvisllnsdgvgrrrlwiglqlepgcgdkrlgrrlgfgwvgnntsys 120
 QY 121 RWAFLDINGAPLIGPLVAVSAEAATPPSPITWEEQCCVKADGFLCEHFPAATCRPLAV 180
 Db 121 rwafldingapligplvavsaeatppspitweeqccvkadgflcehfpaatcrplav 180

QY 181 EPGAAAAAIVSYTGPFFAARGADFOALPVGSSAAVAPLGIOLMCTAPGAVOGHAREAP 240
 Db 181 epgaaaaaivsytgpfeargadfqaalpvgssaaavaplgilolmctappgavoghareap 240
 QY 241 GAMDCEYENGCGEHACNAIPGAPRCQCTPAGALQADGSCSTASATGSCNDLCHEFCYPNP 300
 Db 241 gamdceyengcgchacnaipgaprcqctpagaalqadgscstasatgscndlcefcypnp 300
 QY 301 DQGSYSCMCEGYRLAADORCEDVDPCILEPSPCPQRCVNTGSGEHCCHPYVDLVDG 360
 Db 301 dpgysysscmecegyrlaadqrcevdpcilepsscpcrcvntgsgchchpyvdlvdg 360
 QY 361 ECVEPVDPCEFRANCEYOCPLNQTSTYLCVCAEGFAPLPHBHRQCMFRCNQTACPADCDPN 420
 Db 361 ecvepvdpccfranceygcqplnqtsylcvcaeafiapihpncqmfncqtaacpdcqpn 420
 QY 421 TQASCCEPEGYILDGDFICTDIDECENGCGSCVCHNLDEFTFCICGPPDSALARIHGTDC 480
 Db 421 tqascecepegyilddgfictdidecengcgscvchnlpgtfecicgppdsalarhgtc 480
 QY 481 DSGKVDGSDSGSEPPSPPTGSLTPPAVGLVHS 516
 Db 481 dsgkvdgsgdsgeppspptgsltpavglvhs 516

RESULT 12

AAp22018
 ID AAP22018 standard; Protein; 516 AA.
 AC AAP22018;
 XX
 PD 03-JUL-1992 (first entry)
 DE Human thrombomodulin (1-516) with Asp367 substituted by Glu.
 KW Mutant; thrombin binding site; blood clotting; Tm2 mutator.
 KW Homo sapiens.
 OS
 XX
 PN EP474273-A.
 XX
 PD 11-MAR-1992.
 XX
 PE 05-AUG-1991; 91EP-0202009.
 XX
 PR 03-AUG-1990; 90JP-0204978.
 XX
 PA (ASAHI) ASahi KASEI Kogyo.
 XX
 PI Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;
 XX
 DR WPI: 1992-081820/11.
 XX

New polypeptide inhibits blood coagulation and platelet
 aggregation - promotes thrombin catalysed protein C activation
 for treating myocardial infarction, thrombosis, embolism, etc.
 Example 1; Page 18; 112pp; English.

Plasmid M13TMD1 (see AAP22016) encoding the first 516 N-terminal
 amino acids of human thrombomodulin was used as a template for
 site-directed mutagenesis using the "Tm2 mutator" to produce
 plasmid M13TMM2. In the mutant thrombomodulin encoded by the
 plasmid, the wild-type Asp residue at position 367 is substituted
 by an Glu residue. The activity of this mutant, truncated
 thrombomodulin (i.e. D123Glu) in the activation of protein C was
 compared to that of similarly truncated thrombomodulin with Asp
 at position 367 (i.e. D123Asp). The activity of D123Glu was as high
 as 1.2 times that of D123Asp. See AAP22013-R22022 and AAQ25074.

Sequence 516 AA;

Query Match 99.7%; Score 2908; DB 13; Length 516;
Best Local Similarity 99.6%; Pred. No. 1.3e-162;
Matches 516; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVVLGALALAGLPAPAEPOPGSGQVEHDCFALYPGATFLMASQICDGLRGHLM 60
DB 1 mglvylgalalaglgfpapeepqpgsqvchdcfallypapatflmasqicdglrghlm 60
QY 61 TVRSSVADVLSLLNGSGVGRRLMIGLQLPCCGPKRLGRLRGWVTGDNNTSYS 120
DB 61 tvrsadvlsllngsgvgrrrlmigllqlppcgpkrlgrrlgwvtdnntsys 120
QY 121 RWARLDLNGAPLCGPLCVASAATVPSEPTMEQCEVADGFLCFEHPATCRPLAV 180
DB 121 rwarldlmgaplcgplcvasaaatvpseptmeeqcevadgflcfelhpactrplav 180
QY 181 EPGAANAASITVGPFAARGADFOALPVGSSAAVAPLGLQLMCTAPGAVGHWAREAP 240
DB 181 epgaanaasitvgtpfargadfgalpyvssaaavaplgqlmctappgavghwareap 240
QY 241 GAMDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcehacnaipgarcqcpagallqadgrsctasatqscndlcehfcvnp 300
QY 301 DQPGSYSCMCTGYRLADQHRCEVDDCILPSPCPQRCVNTQGGFECHCYPNYDLYDG 360
DB 301 dqpgyscmctgyrladqhrcevdcdcllpspcprcvntqggfchcypnydlydg 360
QY 361 ECVEPVDFCFRANCEYOCOPINOTSYLCAEGFAPRPHBRCOMFCNOTCAPDCDPN 420
DB 361 ecvepvdfcfranceyocqplnqtsylcvcaeagfaprhbrcqmfncqtacpadcpn 420
QY 421 TQASCECPGYLLDDGFICTDIDECENGFCGVCVCHNLPGTEFCICGPDALARIHIGTDC 480
DB 421 tqascecpgyllddgfictdidceengfcsgvchnlpgtfecicgpdalarihigtcd 480
QY 481 DSGKVDGSDSGEPPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 dsgkvdgsgsgppspptpgstltppavglvhs 516

RESULT 13
AAR31572
ID AAR31572 standard; Protein: 575 AA.
XX
AC AAR31572;
XX
DT 27-MAY-1993 (first entry)
XX
DE Human thrombomodulin.
XX
KW Mutagenesis; site specific; nucleic acid constructs;
XX restriction site; introduction; removal.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT Domain /note= "signal peptide"
FT Domain /note= "EGF-1"
FT Domain 288..323
FT Domain /note= "EGF-2"
FT Domain 329..361
FT Domain /note= "EGF-3"
FT Domain 368..404
FT Domain /note= "EGF-4"
FT Domain 408..439
FT Domain /note= "EGF-5"
FT Domain 445..480
FT Domain /note= "EGF-6"
FT Domain 481..515

FT /note= "O-linked glycosylation domain"
FT Region 516..538
FT /note= "stop transfer sequence"
FT Domain 539..575
FT /note= "cytoplasmic domain"

PD W09301282-A.
XX 21-JAN-1993.
XX
XX 01-JUL-1992; 92WO-0505573.
XX
XX 01-JUL-1991; 91US-0724237.
XX
XX (BERL-) BERLEX LAB INC.
XX
XX Andrews WH, Morser MJ, Vliander LR;
XX
XX WPI: 1993-045488/05.
XX
XX Site-specific mutagenesis of nucleic acid constructs - using an
XX oligo:nucleotide which changes a nucleotide and introduces or
XX removes a restriction site
XX
XX Example: fig 3; 87pp; English.
XX
XX The sequence is that of the native human thrombomodulin showing the
XX six EGF-like domains.
XX
SQ Sequence 575 AA:

Query Match 99.7%; Score 2908; DB 14; Length 575;
Best Local Similarity 99.8%; Pred. No. 1.4e-162;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVVLGALALAGLPAPAEPOPGSGQVEHDCFALYPGATFLMASQICDGLRGHLM 60
DB 1 mglvylgalalaglgfpapeepqpgsqvchdcfallypapatflmasqicdglrghlm 60
QY 61 TVRSSVADVLSLLNGSGVGRRLMIGLQLPCCGDKRRLGRLRGWVTGDNNTSYS 120
DB 61 tvrsadvlsllngdsgvgrrrlmigllqlppcgdkrrlgrrlgwvtdnntsys 120
QY 121 RWARLDLNGAPLCGPLCVASAATVPSEPTMEQCEVADGFLCFEHPATCRPLAV 180
DB 121 rwarldlmgaplcgplcvasaaatvpseptmeeqcevadgflcfelhpactrplav 180
QY 181 EPGAANAASITVGPFAARGADFOALPVGSSAAVAPLGLQLMCTAPGAVGHWAREAP 240
DB 181 epgaanaasitvgtpfargadfgalpyvssaaavaplgqlmctappgavghwareap 240
QY 241 GAMDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcehacnaipgarcqcpagallqadgrsctasatqscndlcehfcvnp 300
QY 301 DQPGSYSCMCTGYRLADQHRCEVDDCILPSPCPQRCVNTQGGFECHCYPNYDLYDG 360
DB 301 dqpgyscmctgyrladqhrcevdcdcllpspcprcvntqggfchcypnydlydg 360
QY 361 ECVEPVDFCFRANCEYOCOPINOTSYLCAEGFAPRPHBRCOMFCNOTCAPDCDPN 420
DB 361 ecvepvdfcfranceyocqplnqtsylcvcaeagfaprhbrcqmfncqtacpadcpn 420
QY 421 TQASCECPGYLLDDGFICTDIDECENGFCGVCVCHNLPGTEFCICGPDALARIHIGTDC 480
DB 421 tqascecpgyllddgfictdidceengfcsgvchnlpgtfecicgpdalarihigtcd 480
QY 481 DSGKVDGSDSGEPPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 dsgkvdgsgsgppspptpgstltppavglvhs 516

[illegible]

| | |
|----|---|
| XX | RESULT 15 |
| XX | AAR20639 |
| XX | ID AAR20639 standard; Protein; 575 AA. |
| XX | AC AAR20639; |
| XX | DT 01-MAY-1992 (first entry) |
| XX | DE Human urinary thrombomodulin. |
| XX | KW Blood clotting; intravascular coagulation. |
| XX | OS Homo sapiens. |
| XX | FH Key Location/Qualifiers |
| XX | FT Peptide 1..18 |
| XX | FT /label= signal |
| XX | FT Protein 19..575 |
| XX | FT /label= thrombomodulin |
| XX | PN MO9200325-A. |
| XX | PD 09-JAN-1992. |
| XX | PE 27-JUN-1991; 91WO-JP00873. |
| XX | PR 27-JUN-1990; 90JP-0168766. |
| XX | PA (MOCH) MOCHIDA PHARM KK. |
| XX | PI Nil A, Morishita H, Uemura A, Mochida E; |
| XX | DR WPJ: 1992-041517/05. |
| XX | N-PSDB; AAQ20810. |
| PT | New modified recombinant human urinary thrombomodulin - has |
| PT | thrombin binding, anticoagulant and thrombolytic activity and is |
| PT | for treating blood coagulation disorders |
| XX | ClaIm 1; Fig 3; 107p; Japanese. |
| CC | A 2.5kb cDNA fragment was isolated by screening a human placental |
| CC | cell library with a probe coding for the N-terminal of human |
| CC | thrombomodulin. This full-length sequence was modified to produce |
| CC | two alternative, truncated forms of the coding sequence encoding |
| CC | only up to amino acid 456 of the mature polypeptide; the Ala |
| CC | residue at position 455 is substituted by Val in one of the two |
| CC | truncated forms. |
| XX | Sequence 575 AA; |
| XQ | |

| | | | | |
|---------------------------|--------|---|-----------|-------------|
| Query Match | 99.6%; | Score 2903; | DB 13; | Length 575; |
| Best Local Similarity | 99.6%; | Pred. No. 2.8e-162; | | |
| Matches 514; Conservative | 0; | Mismatches 2; | Indels 0; | Gaps 0; |
| 0Y | 1 | MIGVLTGALALAGTGGFPAPAEPPGGSGCCVEHDCFALYPGPATLTNASQICDGI | GHLM | 60 |

Db 1 mlgvlylgatlaaglgfpapaepqpgsgqcvchdcfalyppatflinasqldqllrqqlm 60
QY 61 TVRSSVADYISLLNGDGVGRRRLMIGQLPGCCGDPKRLRGLRGFWWTGDNNTSYS 120
Db 61 tvrssvpadvlsllngdgvgrrrllwlgqlppgcgdkrlgplrgfwltdgnntsys 120
QY 121 RMARLDINGAPLCGPLCVANSAEATVPSEPTWEEQCEVAKADGFLCEHFHPATCRPLAV 180
Db 121 rwarldingaplcgplcvavsaatvpsepiweeqqcevkadgflcehfpatcrplav 180
QY 181 EPGAAAAVITGTTPAARGADRFALPVGSAVAPLGLQIMCTAPPGAVQGHWAAREAP 240
Db 181 epgaaaaavsltygtptaargadqlapvgssaavaplgqlmctappgavqghwareap 240
QY 241 GAWDCSVENGCEHACNAIPGAPRCOPAGALQADGRSCTASATOSCNDLCEHRCVNP 300
Db 241 gawdcsvenggcehacnaipgaprcqcpagalaqadgrsctasatqscndlcehrcvnp 300
QY 301 DQPGSTSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTOGGFECHCYPNYDLVDG 360
Db 301 dqpgsyscmctgyrllaadqhrcevdvdcillepspcrqvcvntoggfecchcypnydlvdg 360
QY 361 ECVPEVDPCEFRANCEVOCQPLNQTSTYLCVCAEGFAPTPHEPHRCQMFCNQTACPADCDPN 420
Db 361 ecvpevpdpctranceyqcqplnqtsylcvcaegfaptphephrcqmfcnqtacpadcdpn 420
QY 421 TQASCECEPEGYILDDGFICTDIDECENGSGVCVCHNLPGTFBCTCGPDSALARHIGTDC 480
Db 421 tqascecepegyllddgtfctdidecengsfcsyvcnmlpgtfecicgpdسالارhigtcd 480
QY 481 DSGKVDGDSGSGEPPSPPTPGSTLTPPAVGLVHSG 516
Db 481 dsqkvdgdsqsggeppspptpgstltppavglvhsq 516

Search completed: May 8, 2002, 12:38:54
Job time: 215 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:35:54 ; Search time 29.56 Seconds
(without alignments)
1329.705 Million cell updates/sec

Title: US-09-509-994-1
Perfect score: 2916
Sequence: 1 MGVVLGALALAGLGPAP.....PSPPTGSLTPPAVLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR:68:***
2: PIR:***
3: PIR:***
4: PIR:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------------------|--------------------|
| 1 | 2916 | 100.0 | 575 | 1 THHUB | thrombomodulin pre |
| 2 | 1899 | 65.1 | 577 | 2 A60501 | thrombomodulin pre |
| 3 | 1092.5 | 37.5 | 356 | 2 A25918 | thrombomodulin - b |
| 4 | 416 | 14.3 | 1620 | 2 T27283 | hypothetical prote |
| 5 | 414 | 14.2 | 1574 | 2 T13854 | MESF6 protein - ra |
| 6 | 373 | 12.8 | 2907 | 2 A57278 | fibriillin-2 precu |
| 7 | 367.5 | 12.6 | 1184 | 2 A55184 | fibriillin-2 precu |
| 8 | 367 | 12.6 | 2871 | 2 A55184 | fibriillin-2 precu |
| 9 | 365 | 12.5 | 2918 | 2 A54105 | fibriillin-2 precu |
| 10 | 363 | 12.4 | 3002 | 2 A47221 | fibriillin-1 precu |
| 11 | 361 | 12.4 | 2871 | 2 A55624 | fibriillin-1 precu |
| 12 | 359 | 12.3 | 741 | 2 T46488 | hypothetical prote |
| 13 | 357 | 12.2 | 1221 | 2 A49457 | fibulin-2 precu |
| 14 | 331 | 11.4 | 589 | 2 T43210 | fibulin-1D precu |
| 15 | 329.5 | 11.3 | 1964 | 2 T09059 | notch4 - mouse |
| 16 | 327.5 | 11.2 | 1712 | 2 A38261 | masking protein pr |
| 17 | 324.5 | 11.1 | 689 | 2 T42760 | fibulin, splice fo |
| 18 | 324.5 | 11.1 | 712 | 2 T42990 | fibulin 1, splice |
| 19 | 320.5 | 11.0 | 2321 | 2 S78856 | notch3 protein - h |
| 20 | 320 | 11.0 | 1394 | 2 A35636 | transforming growt |
| 21 | 314.5 | 10.8 | 3507 | 2 T34513 | hypothetical prote |
| 22 | 313 | 10.7 | 601 | 2 B36346 | fibulin 1 precu |
| 23 | 313 | 10.7 | 683 | 2 C36346 | fibulin 1 precu |
| 24 | 313 | 10.7 | 798 | 2 T22793 | hypothetical prote |
| 25 | 312.5 | 10.7 | 1820 | 2 A55494 | latent transformin |
| 26 | 308 | 10.6 | 685 | 2 S78040 | fibulin, splice fo |
| 27 | 308 | 10.6 | 705 | 2 S74968 | fibulin, splice fo |
| 28 | 303.5 | 10.4 | 1251 | 2 A57293 | latent transformin |
| | | | | Notch-1 protein - | |

| | | | | | |
|----|-------|-----|------|----------|---------------------|
| 30 | 286 | 9.8 | 2471 | 2 A49128 | cell-fate determin |
| 31 | 285 | 9.8 | 2352 | 2 T30201 | Notch homolog prot |
| 32 | 283.5 | 9.7 | 2437 | 2 S42612 | transmembrane prot |
| 33 | 280.5 | 9.6 | 2703 | 1 A24420 | notch protein - fr |
| 34 | 280 | 9.6 | 2318 | 2 S45306 | notch 3 protein - |
| 35 | 278.5 | 9.5 | 387 | 2 I38449 | extracellular prote |
| 36 | 277.5 | 9.5 | 511 | 2 T17298 | hypothetical prote |
| 37 | 277 | 9.5 | 2531 | 2 S18188 | notch protein homo |
| 38 | 276.5 | 9.5 | 493 | 2 JC5621 | epidermal growth f |
| 39 | 272 | 9.3 | 2555 | 2 A40043 | notch protein homo |
| 40 | 269.5 | 9.2 | 1217 | 1 EGMSNG | epidermal growth f |
| 41 | 263 | 9.0 | 1203 | 2 A49175 | Notch B protein - |
| 42 | 259.5 | 8.9 | 2531 | 2 T31070 | notch homolog - se |
| 43 | 256.5 | 8.8 | 2524 | 2 A35844 | notch protein - Af |
| 44 | 254.5 | 8.7 | 674 | 2 I55476 | growth potentialin |
| 45 | 249.5 | 8.6 | 1133 | 1 EGRT | epidermal growth f |

ALIGNMENTS

RESULT 1
THHUB
thrombomodulin precursor [validated] - human
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1988 #sequence_revision 12-May-1995 #text_change 15-Sep-2000
C/Accession: A41442; A28307; A29680; A27073; JX0264; 538954
R/Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maru
J. Biochem. 103, 281-285, 1988
A/Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed
A/Reference number: A41442; MUID:88227901
A/Accession: A41442
A/Molecule type: DNA
A/Residues: 1-575 <SH1>
A/Cross-references: DBJ:J02010; MID:9220126; PIDN:BA00149.1; PID:9220127
R/Jackman, R.W.; Beeler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.
Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987
A/Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of t
A/Reference number: A28307; MUID:87317665
A/Accession: A28307
A/Molecule type: DNA
A/Residues: 1-472, 'A', 474-575 <JNC>
A/Cross-references: GB:J02973; MID:9339656; PIDN:AAA61175.1; PID:9339659
R/Suzuki, K.; Kusumoto, H.; Deyashiki, Y.; Nishioke, J.; Maruyama, T.; Kusht, M.; I
EMBO J. 6, 1891-1897, 1987
A/Title: Structure and expression of human thrombomodulin, a thrombin receptor on
A/Reference number: A29680; MUID:88004395
A/Accession: A29680
A/Molecule type: mRNA
A/Residues: 1-575 <SU2>
A/Cross-references: GB:X05495; MID:937123; PIDN:CA29045.1; PID:9736251
A/Experimental source: Lung endothelium
A/Note: part of this sequence, including the amino end of the mature protein, were
R/Wen, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.
Biochemistry 26, 4350-4357, 1987
A/Title: Human thrombomodulin: complete cDNA sequence and chromosome localization
A/Reference number: A27073; MUID:88024950
A/Accession: A27073
A/Molecule type: mRNA
A/Residues: 1-472, 'A', 474-575 <WEN>
A/Cross-references: GB:M16552; MID:9339656; PIDN:AMB59508.1; PID:9339657
A/Experimental source: Placenta
A/Note: parts of this sequence were determined by protein sequencing
R/Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Okuchi, M.; Kimura, S.; Aoki, N.
J. Biochem. 113, 433-440, 1993
A/Title: Urinary thrombomodulin, its isolation and characterization.
A/Reference number: JX0264; MUID:93293792
A/Accession: JX0264
A/Molecule type: protein; mRNA
A/Residues: 19-472, 'A', 474-486 <YAM>
A/Experimental source: urine
A/Note: the urinary form appears to be identical with that circulating in plasma
R/Gellitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell,

Biochem. J. 295, 131-140, 1993
A:Title: Identification of the predominant glycosaminoglycan-attachment site in soluble
serine.
A:Reference number: S38954; MUID:94029900
A:Accession: S38954
A:Molecule type: protein
A:Residues: 475-491; X: 493-494 <GER>
A:Note: The residue designated 'X' was determined to be a Ser with covalently bound choro
R:Meininger, D.P.; Komives, E.A.
Submitted to the Brookhaven Protein Data Bank, September 1995
A:Reference number: A67369; PDB:1ZAO
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
R:Titinsky, A.; Mathews, I.I.
Submitted to the Brookhaven Protein Data Bank, August 1994
A:Reference number: A52804; PDB:1HFT
A:Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442
R:Rabal, R.; Komives, E.A.; Nl, F.
Submitted to the Brookhaven Protein Data Bank, November 1995
A:Reference number: A65583; PDB:1FGD
A:Contents: annotation; conformation by (1)H-NMR, residues 427-444
R:Rabal, R.; Komives, E.A.; Nl, F.
Protein Sci. 5, 195-203, 1996
A:Title: Structural resiliency of an EGF-like subdomain bound to its target protein, the
A:Reference number: A58595; MUID:96276211
A:Contents: annotation; conformation by (1)H-NMR
C:Genetics:
A:Gene: GDB:THBD
A:Cross-references: GDB:119613; OMTM:188040
A:Map position: 20p11.2-20p11.2
A:Intons: #status absent
C:Complex: homodimer, urinary form
C:Function:
A:Description: Inhibits thrombin activation of fibrinogen; cofactor for thrombin activat
A:Pathway: blood coagulation moderation
A:Note: the membrane-bound form is located on the endothelium luminal surface of arterie
A:Note: thrombin complexed with the membrane-bound form is subject to endocytosis
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
C:Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coag
e protein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>
F:19-513/Domain: extracellular #status predicted <EXT>
F:39-486/Product: thrombomodulin, urinary form #status experimental <MAT>
F:24-167/Domain: C-type lectin homology <LCH>
F:17-199/Region: PEST sequence
F:201-233/Region: PEST sequence
F:245-280/Domain: EGF homology <EG1>
F:288-323/Domain: EGF homology <EG2>
F:329-362/Domain: EGF homology <EG3>
F:369-404/Domain: EGF homology <EG4>
F:408-439/Domain: EGF homology <EG5>
F:445-480/Domain: EGF homology <EG6>
F:485-513/Region: PEST sequence
F:517-539/Domain: transmembrane #status predicted <TM>
F:540-575/Domain: intracellular #status predicted <INT>
F:14,115,116,382,409/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:174,225,411,504/Binding site: carbohydrate (Thr) (covalent) #status predicted
F:252-265,267-280,288-296,292-308,310-333,329-340,336-349,351-362,369-378,374-38
F:334,498/Binding site: carbohydrate (Ser) (covalent) #status predicted
F:342/Modified site: erythro-beta-hydroxyaspartic acid (Asn) #status experimental
F:490,492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimen

Query Match 100.0%; Score 2916; DB 1; Length 575;
Best local similarity 100.0%; Pred. No. 9e-169;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVYVIGALALAGLGFAPAPAPORGGSCVHDFALYPPAPFLNAAOICGIGLGHM 60
DB 1 MGVYVIGALALAGLGFAPAPAPORGGSCVHDFALYPPAPFLNAAOICGIGLGHM 60
QY 61 TVSSVAADVISLLNGDGVGRRRLWIGLQIPGCGDPKRLGFLNGFOWTGDNNNTSYS 120
DB 61 TVSSVAADVISLLNGDGVGRRRLWIGLQIPGCGDPKRLGFLNGFOWTGDNNNTSYS 120

DB 61 TVSSVAADVISLLNGDGVGRRRLWIGLQIPGCGDPKRLGFLNGFOWTGDNNNTSYS 120
QY 121 RMRRLNLNGAPLCPCLVAVSAEATVPSPIWEEOCEKAGGFCERFPATCAPLAY 180
DB 121 RMRRLNLNGAPLCPCLVAVSAEATVPSPIWEEOCEKAGGFCERFPATCAPLAY 180
QY 181 EPGAAAAVSTYGPFPAPAGADFOALPVSSAAVAPLGLMCTAPPAVAGHAAEAP 240
DB 181 EPGAAAAVSTYGPFPAPAGADFOALPVSSAAVAPLGLMCTAPPAVAGHAAEAP 240
QY 241 GAMDCEVNGGCEHACNAIPGAPRCQCPAGAAADORSCTASATQSCNDICEHCVPMP 300
DB 241 GAMDCEVNGGCEHACNAIPGAPRCQCPAGAAADORSCTASATQSCNDICEHCVPMP 300
QY 301 DPGSTSCMCEGYRLAADQRCEDVDCILEPSPCQRCVNTQGGFECHECIPNDLVNG 360
DB 301 DPGSTSCMCEGYRLAADQRCEDVDCILEPSPCQRCVNTQGGFECHECIPNDLVNG 360
QY 361 ECEPVPDPFRANCEYOCPLNOTSYLCVCAEGFAPRPHRPHRCOMFCNOTACPADCPFN 420
DB 361 ECEPVPDPFRANCEYOCPLNOTSYLCVCAEGFAPRPHRPHRCOMFCNOTACPADCPFN 420
QY 421 TQASCEPEGYTLDDGFTCTDIDECENGFCGVCNMLPPTFCIGPDSALVRIHTTC 480
DB 421 TQASCEPEGYTLDDGFTCTDIDECENGFCGVCNMLPPTFCIGPDSALVRIHTTC 480
QY 481 DSGKVDGSDSGSEPPSPFPSTLTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSEPPSPFPSTLTPPAVGLVHSG 516

RESULT 2
A60501
thrombomodulin precursor - mouse
N:Alternate names: fetomodulin
C:Species: Mus musculus (house mouse)
C:Date: 17-Apr-1993 #sequence revision 17-Apr-1993 #text-change 16-Jul-1999
C:Accession: S08488; A32001; A60501
R:Dittman, W.A.; Majerus, P.W.
Nucleic Acids Res. 17, 802, 1989
A:Title: Sequence of a cDNA for mouse thrombomodulin and comparison of the predicted
A:Reference number: S08488; MUID:89120454
A:Accession: S08488
A:Molecule type: mRNA
A:Residues: 1-577 <DIT>
A:Cross-references: EMBL:X14432; NID:954781; PIDN:CAA32597.1; PID:954782
R:Dittman, W.A.; Kumada, T.; Sadler, J.E.; Majerus, P.W.
J. Biol. Chem. 263, 15815-15822, 1988
A:Title: The structure and function of mouse thrombomodulin. Phorbol myristate acetat
A:Reference number: A32001; MUID:89008498
A:Accession: A32001
A:Molecule type: mRNA
A:Residues: 97-577 <DIT>
A:Cross-references: GB:J04060
R:Imada, S.; Yamaguchi, R.; Nagumo, M.; Katayanagi, S.; Iwasaki, H.; Imada, M.
Dev. Biol. 140, 113-122, 1990
A:Title: Identification of fetomodulin, a surface marker protein of fetal development
A:Reference number: A60501; MUID:90292331
A:Accession: A60501
A:Molecule type: protein
A:Residues: 19-22,330-343,479-489,545-555,562-575 <IMA>
C:Comment: Thrombomodulin binds to and internalizes with thrombin. It is also a cofac
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology
C:Keywords: anticoagulant; membrane protein; phosphoprotein; receptor
F:24-165/Domain: C-type lectin homology <LCH>
F:244-279/Domain: EGF homology <EG1>
F:287-322/Domain: EGF homology <EG2>
F:328-361/Domain: EGF homology <EG3>
F:368-403/Domain: EGF homology <EG4>
F:407-438/Domain: EGF homology <EG5>
F:444-479/Domain: EGF homology <EG6>

| | | | | |
|-----------------------|------------------|--------------------|------------|--------------|
| Query Match | 14.28; | Score 414; | DB 2; | Length 1574; |
| Best Local Similarity | 34.1%; | Pred. No. 1.2e-17; | | |
| Matches 104; | Conservative 28; | Mismatches 101; | Indels 72; | Gaps 17; |

| | |
|--------|---|
| RESULT | 6 |
| A57278 | |

| | | | | |
|---------------------------|--------|--------------------|------------|--------------|
| Query Match | 12.88; | Score 373; | DB 2; | Length 2907; |
| Best Local Similarity | 32.98; | Pred. No. 5.7e-15; | | |
| Matches 102; Conservative | 24; | Mismatches 92; | Indels 92; | Gaps 18; |

| | | | | |
|----|------|----------|----|------|
| Qy | 474 | RHIG--TD | CD | 481 |
| | | | | |
| | | | | |
| Db | 1518 | RTGGNCTD | TD | 1527 |

A;Cross-references: GDB:293037; OMIM:135821
A;Map position: 3p25-3p24

| | | | | | |
|----|---|--------------|-------------------|-----------------|----------------------|
| | Query March | 12.6% | Score 367.5; | DB 2: | Length 1184; |
| | Best Local Similarity | 30.0% | Pred. No. 5.7e-11 | | |
| | Matches 119; | Conservative | 29; | Mismatches 124; | Indels 125; Gaps 21. |
| Oy | 154 EEOCEYKAP-GGLCFEHPAC-----RPVY-----EEGAAVAAYSLTYGPFRAG | 201 | | | |
| | | | | | |
| Dd | 531 EGQCESNPMLGPCN-IYMISCSBESEPLIVETVRPRPPAARPVYS-----EAEM | 582 | | | |
| Oy | 202 ADFOALEYVGSSAAV---APLGTLQAKTPPAGVAGCHMAREARPGAMDCSVEGGCGHRCAN | 258 | | | |
| | | | | | |
| Dd | 583 AGRAALSGLEAEIPNSTLEDODDECLLPGEL-----CQHLCIN | 622 | | | |


```

QY      259  IPGAPRCCRGALADQDGRSC-----TASAT-----QS 287
Db      623  TVGSHYHACAPGGFSLQDDGRCRPEGHPPQEPAPQEPALAKSEFSQVAVNTLPPLPQDNT 682
QY      288  CND--LCEHFCVPRPDQPGSYSCMCETGYRLAADQHRCEVDQDCLLEBSPCP--QRCVNT 343
Db      683  CKDQGPCKQVC---STVGSAICSCFPETAYIMADGVSCEDINECYTDLHTSGRGEHYNT 739
QY      344  QGGEFCH-----CYPNYDLVDGCEVEPYDPCEFRANCXYOCPLNQTSLCV----- 389
Db      740  LGSHTCYKALTCCEGYALKDQEC--EDVDEC--AMGTHTCQP----GFLQNMTKGSFYQQA 792
QY      390  --CAEGFAPRIPH-----EPHRCMFCNQTAQCPADCPNTQASCECEPGY-I 432
Db      793  RQRCHMDGFLQDPQEGNCVDINECTSLSEPCRPFGSCINTGVSYTCQNPILI--CARGYHA 849
QY      433  LDQGFICTDIDECENGCFCSG---VCHNLPGTFECIC 466
Db      850  SDDGAKCVDVNECEETGVHRCGEGQVCHNLPGSYRDCD 886

RESULT      8
A55567
C: fibrillin I - bovine
C: Species: Bos primigenius taurus (cattle)
C: Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 11-Jan-2000
C: Accession: A55567
C: R: Ilistra, D.J.: L1, L.; Potter, K.A.; Womack, J.; Byers, P.H.
Genomics 23, 480-485, 1994
A: Title: Sequence of the coding region of the bovine fibrillin cDNA and localization to
A: Reference number: A55567; MUID: 95137597
A: Accession: A55567
A: Status: Preliminary
A: Molecule type: mRNA
A: Residues: 1-2871 <NTL>
A: Cross-references: GB:I8748; NID: g508427; PIDN: AAA74122.1; PID: g508428
C: Superfamily: unassigned EGF-related proteins; EGF homology
F: I1201-1236/Domain: EGF homology <EGF>

```

```

Query Match Similarity      12.6%:  Score 367;  DB 2;  Length 2871;
Best Local Similarity      31.5%:  Pred. No.1.3e-14;
Matches 111;  Conservative 26;  Mismatches 109;  Indels 106;  Gaps 20.

Oy  244  DCSVNGGCEHNCNMPGAPRCQCPAGALQADGRCSTASATQSCNDLCEHRCVNP--- 300
      :||: |||||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db  1200  EESIMNGGCEFTCTNSSESYECSQCGPFLAMPDQRCT-----DIDE--CEDPNPIC 1249

Oy  301  -----DQPGSYSCMCETGYRLAADQHRCEVDYDCLTEPSPC-PQRCVNTQGGFECHCYP 353
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db  1250  DGGQCTNINIGFERCLCYDGFPMASEDKTKTVDNEDCLNPNICLISGTCENTKGSFICHQM 1309

Oy  354  NNDLVNDE--CYEPVDPCC--FRANCEYQCGPRLNQ--SYLCAQAEFGA-----PIP 398
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db  1310  GYSGKKGKGGCYD-INECETGAHNCRHAHVCTNTAGSFCSGSPGIVGDIKCTDIDEGS 1368

Oy  399  HEPRRCQMCFCNQTACPADCDPNTQAS--CECPRGYILIDBGLICTDIDEC-EN----- 447
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db  1369  NGTHMKSQH-----ADC-KNTMGSYRLCKRKRGY-TGGGFTCTDIDDECESENILCGNGQ 1419

Oy  448  -----GGF-----CS-----GVCNLPDTEPCTIGPPDSALY 473
      :||: |||||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db  1420  CLNAPGGRCECDMGFVPSADGRACEDIDECSPINLCVFETGCHNLPGLFRCCEIEGYELD 1479

Oy  474  RHIG-----TDCDSGK-VDGDSSSGEPSPPTGSLTPRAVLY 513
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db  1480  RSGGNCITVNECLDPTTCLISGNCVNPPTGSYTCQCPD---FELNPTRVACV 1527

RESULT      9
f1dbf1lin-2 precursor - human

```

C:Species: Homo sapiens (man)
C:Date: 09-Sep-1994 #sequence revision 09-Sep-1994 #text_change 17-Nov-2000
C:Accession: A54105; S17063; S31101
R:Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sanguineti, C.; Bonadio, J.; Mecha
R:Cell Biol. 124, 885-883, 1994
A:title: Structure and expression of fibrillin-2, a novel microfibrillar component pr
A:Reference number: A54105; MUID:94165150
A:Accession: A54105
A:Status: Preliminary: nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-2918 <ZHA>
A:Cross-references: GB:U03272
R:Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras,
Nature 332, 330-334, 1991
A:title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff
A:Reference number: S17062; MUID:91304567
A:Accession: S17063
A:Molecule type: mRNA
A:Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLT', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>
A:Cross-references: EMBL:X62009
R:Miliewicz, D.M.
submitted to the EMBL Data Library, December 1992
A:Reference number: S31101
A:Accession: S31101
A:Molecule type: mRNA
A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLT', 1922-1923, 'LD', 1926, 'P',
A:Cross-references: EMBL:X62009
C:Genetics:
A:Gene: GDB:FBN2
A:Cross-references: GDB:128122; OMIM:121050
A:Map position: 5q23-5q31
C:Superfamily: Unassigned EGF-related proteins; EGF homology
C:Keywords: extracellular protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-2918/Product: fibrillin-2 #status predicted <MAT>
F:1245-1280/Domain: EGF homology <EGF1>
F:1970-2013/Domain: EGF homology <EGF>

```

Query Match      12.5%: Score 365; DB 2; Length 2918;
Best Local Similarity 33.1%: Pred. No. 1,7e-14;
Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps 17;

Oy      244 DCSVENGGCEHMACNIPGAPRCQCPAGALQADGRSCTASATQSCNDLCENHFCVPPNP-- 301
          : : ||||| : : : ||||| : : : ||||| : : : |||||
Db      1244 ECMIMNGCDTQCTTNSGSEYSCSGEVALMPGRSCA-----DIDE--CENNPDTIC 1293

Oy      302 -----QPSYSCMCEETGVRLLADQHRREDVDDCLTLESPPC-QRCVNTQSGFECHCYP 353
          : : ||||| : : : ||||| : : : ||||| : : : |||||
Db      1294 DGGCCTNIPGEYRCCLCYDGFMSAMDKCTCIDVNECDLNSNICMFGCECENTKSFICHQQL 1353

Oy      354 NYDLVDE--CVPPVPPC--FRANCEYOCPLN-QTSYLVCACGAPAPPH-----EPHR 403
          : : : ||||| : : : ||||| : : : ||||| : : : |||||
Db      1354 GYSKKKTTCTCTD-VDECELGAINCMHMASCLNIPSPKSCREGV--IGNGIKCIDLDE 1410

Oy      404 COMFCNOTACPADCDPTQAS--CEPEEGTILDDGFLCTDIDE-----CENG----- 448
          : : : ||||| : : : ||||| : : : ||||| : : : |||||
Db      1411 CSNGTHGCSINAOQ-VNTPGSPYRCACSEGF-TGDEGFLCSDVDECAENINLCENGCLNVP 1468

Oy      449 -----GF-----CSGVCHNLPCTFEFCICGDSALVHNIG- 477
          : : ||||| : : : ||||| : : : ||||| : : : |||||
Db      1469 GAYNCECEMGFTTPASDSRSQDIDDECSFONICVSGTCNNLPGMFHCICDGGYELDPTGGN 1528

Oy      478 -TPCD 481
          : : ||
Db      1529 CTDTD 1533

RESULT 10
A47221
fibrillin 1 precursor - human (fragment)
C:Species: Homo sapiens (man)
C:date: 02-Jun-1995 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000

```

C:Accession: A47221; I54355; S17064; I59574; S17062; S62111; A34198
 R:Corson, G.M.; Chabberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.
 Genomics 17, 476-484, 1993
 A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure
 A:Reference number: A47221; MUID:94010947
 A:Accession: A47221
 A:Molecule type: mRNA
 A:Residues: 1-337, 'T', 339-1029 <COR>
 A:Cross-references: GB:X63556
 R:Perella, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Pangillan, T.; Bonad
 Hum. Mol. Genet. 2, 961-968, 1993
 A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene F
 A:Reference number: I54355; MUID:93372860
 A:Accession: I54355
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 132-3002 <PER>
 A:Cross-references: GB:L13923; NID:9306745; PIDN:AA02036.1; PID:9306746
 R:Maslen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.
 Nature 352, 334-337, 1991
 A:Title: Partial sequence of a candidate gene for the Marfan syndrome.
 A:Reference number: S17064; MUID:91304568
 A:Accession: S17064
 A:Molecule type: mRNA
 A:Residues: 1030-3002 <MAS>
 A:Cross-references: EMBL:X63556
 R:Dietz, H.C.; Valle, D.; Franccomano, C.A.; Kendzior, R.J.
 Science 259, 680-683, 1993
 A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.
 A:Reference number: I59574; MUID:93157831
 A:Accession: I59574
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 2217-2288, 'T', 2290-2325 <RES>
 A:Cross-references: GB:S54426; NID:9264860; PIDN:AA025244.1; PID:9264861
 R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattei, M.G.; Sartarazi, M.; Tsipouras, P.
 Nature 352, 330-334, 1991
 A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differe
 A:Reference number: S17062; MUID:91304567
 A:Accession: S17062
 A:Molecule type: mRNA
 A:Residues: 'VATVVFILSTNKM', 944-1444 <LEB1>
 A:Cross-references: EMBL:X62008; NID:913198; PIDN:CAB56534.1; PID:95924015
 A:Accession: S62111
 A:Molecule type: protein
 A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEB2>
 R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.
 J. Biol. Chem. 264, 21381-21385, 1989
 A:Title: Connective tissue microfibrils. Isolation and characterization of three large F
 A:Reference number: A34198; MUID:90078246
 A:Accession: A34198
 A:Molecule type: protein
 A:Residues: 565-575; 1890-1892, 'T', 1894-1900 <MAD>
 C:Comment: Fibrillin is a major component of elastin-associated microfibrils.
 C:Genetics:
 A:Gene: FBN1
 A:Cross-references: GDB:127115; OMIM:134797; OMIM:154700
 A:Map position: 15q21.1-15q21.1
 A:Introns: 2236/1; 2258/1; 2297/1
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein; N
 F:1-3002/Product: fibrillin (5'-region exon A splice form) #status predicted <MTC>
 F:132-3002/Product: fibrillin (5'-region exon C splice form) #status predicted <MTC>
 F:1332-1367/Domain: EGF homology <EGF2>
 F:1457-1492/Domain: EGF homology <EGF2>
 F:2262-2295/Domain: EGF homology <EGF1>

Query Match 12.4%; Score 363; DB 2; Length 3002;
 Best Local Similarity 30.8%; Pred. No. 2,4e-14;
 Matches 111; Conservative 22; Mismatches 105; Indels 122; Gaps 20;

244 DCSVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP--- 300

```

Db 1331 ECSTIMMGCTFCTNBEGBSECCSGPFFALMPDRCT-----DIE--CEDNPNC 1380
Qy 301 -----DQPSYSCMETGYRLAADHRCEDVDICLEBSPC-PQRCVNTGGFEFCHCP 353
Db 1381 DGGGCTNIPGEYRCLCYDGFMASEDMKTCVDVNECDLNNICISGTCENTKGSFICHCMD 1440
Qy 354 NYDLVGE--CVPVDPCC--FRANCYQCPINOT--SYLCVABEFA-----PIP 398
Db 1441 GYSGKKKTKCTD-INECEIGAHNCKHAVCTNTAGSFKSCSPGWIIGDKICTDIDEC 1499
Qy 399 HEPRRCMFQNCOTACPADCPNTQAS--CEPEGYILDDGFTCTDIDEC-EN----- 447
Db 1500 NGTHMCSQH-----ADC-KNTMSYRCLCKEGY-TGGGFTCTDIDECSEMLNLCNGQ 1550
Qy 448 -----GGF-----CS-----GVCHNLPTECTICGPDALV 473
Db 1551 CLNAPGGYRCCECDMGFPASADGRACEDIDECSPNICVGTCHNLPGLERCECEIGYEID 1610
Qy 474 RHIG-----TDGSGKVDGDSGSGEPSPPTGSG-----TLTPPAVLV 513
Db 1611 RSGGNCCTVNECIDPTTCTISGNCVN-----TFESYICDPPPELNPTRVGCY 1658

RESULT 11
A55624
fibrillin-1 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 11-Jan-2000
C:Accession: A55624
R:Yin, W.; Smiley, E.; Germiller, J.; Sanginetti, C.; Lawton, T.; Perella, L.; Ramire
J. Biol. Chem. 270, 1798-1806, 1995
A:Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin
A:Reference number: A55624; MUID:95130561
A:Accession: A55624
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2871 <YIN>
A:Cross-references: GB:I29454; NID:9575509; PIDN:AAA6840.1; PID:9575510
C:Genetics:  

A:Gene: FBN-1  

A:Superfamily: unassigned EGF-related proteins; EGF homology
F:1201-1236/Domain: EGF homology <EGF>
```

Query Match 12.4%; Score 361; DB 2; Length 2871;
 Best Local Similarity 29.4%; Pred. No. 3e-14;
 Matches 117; Conservative 26; Mismatches 127; Indels 128; Gaps 21;

212 SAAVAPRLG-----QLMCTAPPAYOGHWAAREAPGAMDCSVENGCEHACNAIPGAPRC 265
 Db 1162 SANLCPRHRCVNLIGKYQACNPGYHPTDLRCVDIDECISMGCECFCTNSDGSYC 1221
 Qy 266 QCPAGALQADGRSCTASATQSCNDLCEHFCVNP-----DQPSYSCMETGYRL 316
 Db 1222 SCQGFALMPDRCT-----DIDQ--CEDNPNICDGGCTNIPGEYRCLCYDGFMA 1271
 Qy 317 AADHRCEDVDCTLEBSPC-PQRCVNTGGFECHCPYNDLVGE--CVPVDPCC--PR 371
 Db 1272 SEDMKTCVDVNECDLNNITLSTGTCENTKGSFICHCMDGSGKKGCTCTD-INECEITA 1330
 Qy 372 ANCEYQCPINOT--SYLCVABEFA-----DIPHRRCMFQNCOTACPADCPN 420
 Db 1331 HNCGRHNVGNTAGSFKSCSPGWIIGDKICTDLDSCSNTHMCSQH-----ADC-KN 1382
 Qy 421 TQAS--CEPEGYILDDGFTCTDIDEC-EN-----GGF----- 450
 Db 1383 TMSYRCLCKDGY-TGGGFTCTDIDECSEMLNLCNGQCLNAPGGYRCCECDMGFPASADG 1441
 Qy 451 -----CS-----GVCHNLPTECTICGPDALVRIHIG-----TDGSGK 484
 Db 1442 KACEDIDECSLPNICVFGTCHNLPGLERCEIGYELDRSGGNCCTVNECIDPTTCTISGN 1501

QY 485 VDGSGSGGPPSPSPGS-----TLPPAVGLV 513
 Db 1502 CVN-----TPGSTYCCDCSPDFELNPRVGCY 1527

RESULT 12

hypothetical protein DKFZp434J065.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46488

R:Diesterhoef, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: 223035

A:Accession: T46488

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-741 <AAA>

A:Cross-references: EMBL:AL137638

A:Experimental source: adult testis; clone DKFZp434J065

C:Genetics:

A:Note: DKFZp434J065.1

Query Match

Best Local Similarity 12.3%; Score 359; DB 2; Length 741;
 Pred. No. 1,2e-14;

Matches 86; Conservative 41; Mismatches 101; Indels 44; Gaps 15;

QY 245 CSVNGCENACNAIPGAPRCQCPAGALQADGRSCTA-SATQSCNDLCEHFCVPPNDP 303
 Db 68 CAMBDHNCCEOLCVNPPSCFYCCGALAEGRCAVAVYCASENHGCHECV-NAD-- 124
 QY 304 GSYSCMCETGYRLAADHRCEDVDICLIEPSPCPCRCVTOGGFECHCYENYD-VNGEC 362
 Db 125 GSYLCQCHGEFALNPDECTCKIDYCASNRHGCHECVNTDDSYSCGLGFTLNPKKT 184
 QY 363 VEPVDPCE--RANCEYOCQPLNOSTYLCVCAEGFAPRPH-----EPHRCQMEC 408
 Db 185 CRRINCYCMLNKPCCHEHCVNME--SYVCRHRYGTYLDPNGKTCGRVDCADQDHCEBOLC 243
 QY 409 NOTACRPNACDNPQAS--CECPGTYIIDDFG-ICTDIDEC--ENGFGSGVCNHLPGTF 462
 Db 244 -----LNTEDSYVQCCESEGLINEDLKTCSRDVCLLSDHG--CEYSCVNMDSRF 291
 QY 463 ECICGPDALVRIIGTDCDSGRVGD---GDSG 491
 Db 292 ACQC-PEGHVLRSQDKTC--AKLDSICALGDHG 320

RESULT 13

fibulin-2 precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 11-Jan-2000

C:Accession: A49457; S74095

R:Pan, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.

J. Cell Biol. 123, 1269-1277, 1993

A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein with

A:Reference number: A49457; MUID:94064787

A:Accession: A49457

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1221 <PAN>

A:Cross-references: GB:X75285; NID:9437046; PIDN:CAA53040.1; PID:9437047

R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.

Eur. J. Biochem. 240, 427-434, 1996

A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix met

A:Reference number: S74095

A:Accession: S74095

A:Molecule type: protein

A:Residues: 236-238, 'X', 240-247, 260-275, 336-344, 'L', 346-361, 405-426, 566-568, 'EW', 569-589

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer

F:942-978/Domain: EGF homology <EGF>

Query Match 12.2%; Score 357; DB 2; Length 1221;
 Best Local Similarity 29.5%; Pred. No. 2.5e-14;
 Matches 123; Conservative 39; Mismatches 141; Indels 114; Gaps 27;

QY 154 EEOCEVKAAD-GELEFHEPRATC-----RPLAV-----EFGAAAAAS-----IT 192
 Db 521 EEOCEBSPNLTGYPCH-NHMLSCCEGEPLVPEVRRPEBPAPRVSEMAAREALS 579
 QY 193 YGT----PFAAGADPO---ALP-----VGSSAAVAPLQILM-----CTAPPG 229
 Db 580 LGTEALEPNSLFGDDODECLMLPGLCQHLICINTVGYRCAFCFPELQGGGRICRPDRG 639
 QY 230 AVQGHARE-ARGANDCSY-----ENGCEHACNAITGARCCCPAGA 272
 Db 640 APQLDTAREASAPRSASQVSPNTIDLPVQPNPTCKDNPCRQVCRVVDGTAMCSCFPGYA 699
 QY 273 LDADGRSC-----TASAFQSCNDLCEHFCVPPNDPQGSYSYSC-----MCETGYRLAADHR 322
 Db 700 IMADGYSCEDDOECIMGTHDCS--WKQFCV---NTLGSFYCVNHTVLCABEYILNA-HRK 753
 QY 323 CEDVDICLIEPSPC--PQRCVNTQGGFECH---CYPNYDVLVDGCEVPPVDPCEPRA--NC 374
 Db 754 CVDINCEVTDLHTCTRAEHCVNTPGSCFYKALTCFPGVLTGBCSD-VDECVGTGTHNC 812
 QY 375 E--YOCQPLNOSTYLCV---CAGGFAPRPH-----EPHRCQFCNQTACRA 415
 Db 813 QAGFSCQ--TKGSFYCQARQCMDFLQDPGNCVDINECTSLLEPCSSGFSCTNTVGSY 871
 QY 416 DCDPNTQASCEPREGY-IIDDFICTDIDECENGFGSCG---VCHNLPGTECTICGP 468
 Db 872 TQGRNPLV---GGRGHANEBSSECVYDVNECTGVNHRGEGQLCVNLGSTRQCKP 925

RESULT 14

fibulin-1D precursor - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000

C:Accession: T43210

R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.

submitted to the EMBL Data Library, June 1998

A:Description: Identification of chicken and C. elegans fibulin-1 homologs and charac

A:Reference number: 223337

A:Accession: T43210

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-589 <BAR>

A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1

C:Genetics:

A:Note: Intron positions not resolved (incomplete sequence)

C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 11.4%; Score 331; DB 2; Length 589;
 Best Local Similarity 28.1%; Pred. No. 5e-13;
 Matches 126; Conservative 49; Mismatches 166; Indels 108; Gaps 28;

QY 103 GPLRGF-QWVTGDNNTSYRRMARLDLNGAPRLCGPLCVASAAEAVVPSPI-----WEQ 156
 Db 19 GCLRSFNKCCNGDLEITH--ASEIITGRLNDPHVLIHGDRCASSHEHLCNDRGSKV 75
 QY 157 QCEVKA-----DGFLCEHFNP--ATCRPLAVE-----PGAAAASVITGTFPA--- 198
 Db 76 ECCCRSGFLDAPGMACVNHIDECNLMDDCLESQRLNTPSPFCIRTLSCGTGYAADS 135
 QY 199 --ARGADFQALPVGSSAAVAPLGLDMLCTAPPAVQ-----GHVAREAPGAMDCS---- 246
 Db 136 ETERCRDVIDECNLGSH---DCGPLYOCRNTOGSRCAKCKGDELDLNPMTGECTSYTC 191
 QY 247 -----VENGGCE-----HACNA-----IPGAPRQ-----CPAGALQADGRSCTA 282

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:39:20 ; Search time 21.32 seconds
(without alignments)
544,639 Million cell updates/sec

Title: US-09-509-994-2
Perfect score: 2916
Sequence: 1 MLGVLVLGALAGLGPAP.....PSPTPGSLTPPAGLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------------------------|
| 1 | 2916 | 100.0 | 575 | 1 | US-08-312-870-1 |
| 2 | 2916 | 100.0 | 575 | 6 | Sequence 1, Appli Patent No. 5466668 |
| 3 | 2912 | 99.9 | 575 | 1 | US-08-261-206A-59 |
| 4 | 2908 | 99.7 | 575 | 1 | Sequence 59, Appli Patent No. 5256770 |
| 5 | 2846.5 | 97.6 | 572 | 6 | Sequence 54, Appli Patent No. 5256770 |
| 6 | 2826 | 96.9 | 498 | 2 | Sequence 2, Appli Patent No. 5256770 |
| 7 | 2824 | 96.8 | 497 | 1 | Sequence 3, Appli Patent No. 5256770 |
| 8 | 2766 | 94.9 | 494 | 1 | Sequence 14, Appli Patent No. 5256770 |
| 9 | 2766 | 94.9 | 494 | 1 | Sequence 14, Appli Patent No. 5256770 |
| 10 | 2764 | 94.8 | 494 | 1 | Sequence 14, Appli Patent No. 5256770 |
| 11 | 2764 | 94.8 | 494 | 1 | Sequence 14, Appli Patent No. 5256770 |
| 12 | 2690 | 92.2 | 475 | 1 | US-08-110-011A-16 |
| 13 | 2690 | 92.2 | 475 | 1 | Sequence 2, Appli Patent No. 5256770 |
| 14 | 2686 | 92.1 | 475 | 1 | Sequence 2, Appli Patent No. 5256770 |
| 15 | 2686 | 92.1 | 475 | 1 | Sequence 2, Appli Patent No. 5256770 |
| 16 | 2680 | 91.9 | 476 | 1 | US-08-587-389-1 |
| 17 | 2680 | 91.9 | 476 | 1 | Sequence 1, Appli Patent No. 5256770 |
| 18 | 2678 | 91.8 | 476 | 1 | Sequence 1, Appli Patent No. 5256770 |
| 19 | 2678 | 91.8 | 476 | 1 | Sequence 2, Appli Patent No. 5256770 |
| 20 | 2678 | 91.8 | 476 | 1 | Sequence 18, Appli Patent No. 5256770 |
| 21 | 2678 | 91.8 | 476 | 1 | Sequence 2, Appli Patent No. 5256770 |
| 22 | 2596 | 89.0 | 456 | 1 | US-08-110-011A-18 |
| 23 | 2596 | 89.0 | 456 | 1 | Sequence 18, Appli Patent No. 5256770 |
| 24 | 2592 | 88.9 | 456 | 1 | Sequence 4, Appli Patent No. 5256770 |
| 25 | 2592 | 88.9 | 456 | 1 | Sequence 4, Appli Patent No. 5256770 |
| 26 | 2543 | 87.2 | 446 | 1 | US-08-587-389-3 |
| 27 | 2543 | 87.2 | 446 | 1 | Sequence 5, Appli Patent No. 5256770 |

| | | | | | | |
|----|-------|------|------|---|-------------------|--------------------|
| 28 | 1621 | 55.6 | 275 | 1 | US-08-312-870-7 | Sequence 7, Appli |
| 29 | 1159 | 39.7 | 215 | 1 | US-08-312-870-5 | Sequence 5, Appli |
| 30 | 689 | 23.6 | 115 | 1 | US-08-312-870-9 | Sequence 9, Appli |
| 31 | 681 | 23.4 | 114 | 2 | US-08-733-564-1 | Sequence 1, Appli |
| 32 | 584 | 20.0 | 652 | 2 | US-08-751-305-2 | Sequence 2, Appli |
| 33 | 356 | 12.2 | 638 | 2 | US-08-897-443-1 | Sequence 1, Appli |
| 34 | 352 | 12.1 | 58 | 1 | US-08-261-206A-3 | Sequence 3, Appli |
| 35 | 326.5 | 11.2 | 956 | 2 | US-08-897-443-3 | Sequence 3, Appli |
| 36 | 323 | 11.1 | 1233 | 3 | US-08-479-722B-4 | Sequence 4, Appli |
| 37 | 322.5 | 11.1 | 1394 | 6 | 5177197-30 | Patent No. 5177197 |
| 38 | 319.5 | 11.0 | 1833 | 3 | US-08-479-722B-2 | Sequence 2, Appli |
| 39 | 319.5 | 11.0 | 1833 | 5 | PCT-US95-02251-18 | Sequence 18, Appli |
| 40 | 315.5 | 10.8 | 443 | 2 | US-08-833-963C-2 | Sequence 2, Appli |
| 41 | 315.5 | 10.8 | 443 | 3 | US-08-980-514-1 | Sequence 1, Appli |
| 42 | 308 | 10.6 | 448 | 2 | US-08-884-072-1 | Sequence 1, Appli |
| 43 | 308 | 10.6 | 448 | 4 | US-09-212-168-1 | Sequence 1, Appli |
| 44 | 303.5 | 10.4 | 1251 | 5 | PCT-US95-02251-3 | Sequence 3, Appli |
| 45 | 303.5 | 10.4 | 1252 | 1 | US-08-199-780-3 | Sequence 3, Appli |

ALIGNMENTS

RESULT 1
US-08-312-870-1
; Sequence 1, Application US/08312870
; Patent No. 5639625
; GENERAL INFORMATION:
; APPLICANT: Carson, Craig W.
; APPLICANT: Esmon, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to
; TITLE OF INVENTION: Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Richards, Medlock & Andrews
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: Texas
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,870
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hansen, Eugenia S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: OMRF B35150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; NAME/KEY: Protein
; LOCATION: 19..575
US-08-312-870-1

Query Match 100.0%; Score 2916; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 6.9e-199;

| | Matches | 516; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
|----|---------|--------|-------------------|-----------------|-----------------|----------------|-------------|-----|------|----|
| QY | 1 | MLGV | VLGALALAGLGPAPAE | PQPGSQSCVEHDC | FALYPGPATFLNASQ | ICDGLRGHLM | 60 | | | |
| Db | 1 | MLGV | VLGALALAGLGPAPAE | PQPGSQSCVEHDC | FALYPGPATFLNASQ | ICDGLRGHLM | 60 | | | |
| QY | 61 | TVRSS | AAADVISLLLLNGDGGV | RRRLWIGLQLPGCGD | PKRLGRLGFWWTGD | NNNTSYS | 120 | | | |
| Db | 61 | TVRSS | AAADVISLLLLNGDGGV | RRRLWIGLQLPGCGD | PKRLGRLGFWWTGD | NNNTSYS | 120 | | | |
| QY | 121 | RWARI | DLNGAPLCGPCLCVAVS | AAETVPSEPIWEEQ | CEVKADGFLCEFFHP | ATCRPLAV | 180 | | | |
| Db | 121 | RWARI | DLNGAPLCGPCLCVAVS | AAETVPSEPIWEEQ | CEVKADGFLCEFFHP | ATCRPLAV | 180 | | | |
| QY | 181 | EPGAAA | AAVSITVTGTPFAARG | ADFOALPVGSSAAV | PLGLQLMCTAPP | GAVOGHWAREAP | 240 | | | |
| Db | 181 | EPGAAA | AAVSITVTGTPFAARG | ADFOALPVGSSAAV | PLGLQLMCTAPP | GAVOGHWAREAP | 240 | | | |
| QY | 241 | GAWD | CSVENGGCEHACNAIP | GAAPRCQCPAGAAL | QADGRSCTASATQ | SCNDLCEHFCVPNP | 300 | | | |
| Db | 241 | GAWD | CSVENGGCEHACNAIP | GAAPRCQCPAGAAL | QADGRSCTASATQ | SCNDLCEHFCVPNP | 300 | | | |
| QY | 301 | DQPG | SYSCMCETGYRLAAD | QHRCEDVDDCILEP | SPCQRCVNTQGG | FECHCYPNYDLVDG | 360 | | | |
| Db | 301 | DQPG | SYSCMCETGYRLAAD | QHRCEDVDDCILEP | SPCQRCVNTQGG | FECHCYPNYDLVDG | 360 | | | |
| QY | 361 | ECVE | PVDPCFRANCEYQCQ | PLNOTSYLCVCAEG | FAPIPHEPHRCQ | MCFCNTACPADCPN | 420 | | | |
| Db | 361 | ECVE | PVDPCFRANCEYQCQ | PLNOTSYLCVCAEG | FAPIPHEPHRCQ | MCFCNTACPADCPN | 420 | | | |
| QY | 421 | TQAS | CECEPEGYILDDG | FICTDIDECENG | GGFCSGVCHNLPGT | FEICGPD | SALARHIGTDC | 480 | | |
| Db | 421 | TQAS | CECEPEGYILDDG | FICTDIDECENG | GGFCSGVCHNLPGT | FEICGPD | SALARHIGTDC | 480 | | |
| QY | 481 | DSKG | VDGGDGGSGGSEPP | SPPTGTLTPPAY | GLVHSG | 516 | | | | |
| Db | 481 | DSKG | VDGGDGGSGGSEPP | SPPTGTLTPPAY | GLVHSG | 516 | | | | |

RESULT 2
5466668-6
; Patent No. 5466668
; APPLICANT: GLASER, CHARLES B.; MORSER, MICHAEL J.; LIGHT,
; DAVID R.
; TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR
; PHARMACEUTICAL USE
; NUMBER OF SEQUENCES: 57
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/155,346
; FILING DATE: 22-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 568,456
; FILING DATE: 15-AUG-1990
; APPLICATION NUMBER: 506,325
; FILING DATE: 09-APR-1990
; APPLICATION NUMBER: 406,941
; FILING DATE: 13-SEP-1989
; APPLICATION NUMBER: 345,374
; FILING DATE: 28-APR-1989
; SEQ ID NO: 6:
; LENGTH: 575
5466668-6

```
Query Match      100.0%; Score 2916; DB 6; Length 575;
Best Local Similarity 100.0%; Pred. No. 6.9e-199;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--------|-----|------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|---|---|---|---|---|---|---|---|---|---|-----|---|---|---|---|---|---|-----|-----|-----|---|---|-----|-----|
| Db | 61 | TVRSSYAADV | I | S | L | L | N | G | D | G | V | G | R | R | L | W | I | G | L | P | P | C | G | D | P | K | R | L | G | P | F | Q | W | T | G | D | N | T | S | Y | S | 120 | | | | | | | | | | | | | |
| QY | 121 | RWARDLNGAP | L | C | G | P | L | C | V | A | V | S | A | E | A | T | V | P | S | E | T | W | E | B | Q | O | C | E | V | K | A | D | G | F | L | C | E | H | R | P | A | T | C | R | P | L | A | V | 180 | | | | | | |
| Db | 121 | RWARDLNGAP | L | C | G | P | L | C | V | A | V | S | A | E | A | T | V | P | S | E | T | W | E | B | Q | O | C | E | V | K | A | D | G | F | L | C | E | H | R | P | A | T | C | R | P | L | A | V | 180 | | | | | | |
| QY | 181 | EPGAAAAA | V | S | I | T | Y | G | T | P | F | A | A | R | G | A | D | F | Q | A | L | P | V | G | S | S | A | A | V | A | P | L | G | L | Q | L | M | C | T | A | P | P | G | A | V | O | G | H | W | A | R | E | A | P | 240 |
| Db | 181 | EPGAAAAA | V | S | I | T | Y | G | T | P | F | A | A | R | G | A | D | F | Q | A | L | P | V | G | S | S | A | A | V | A | P | L | G | L | Q | L | M | C | T | A | P | P | G | A | V | O | G | H | W | A | R | E | A | P | 240 |
| QY | 241 | GAWDCSVENG | G | C | E | H | A | C | N | A | I | P | G | A | R | C | O | C | P | A | G | A | L | Q | A | D | G | R | S | C | T | A | S | A | T | O | S | C | N | D | L | C | E | H | F | C | V | P | N | P | 300 | | | | |
| Db | 241 | GAWDCSVENG | G | C | E | H | A | C | N | A | I | P | G | A | R | C | O | C | P | A | G | A | L | Q | A | D | G | R | S | C | T | A | S | A | T | O | S | C | N | D | L | C | E | H | F | C | V | P | N | P | 300 | | | | |
| QY | 301 | DQPGSYSCMC | T | G | Y | R | L | A | O | D | H | R | C | E | D | V | D | D | C | I | L | B | P | S | P | C | P | Q | R | C | V | N | T | Q | G | S | F | E | C | H | C | Y | P | N | D | L | V | D | G | 360 | | | | | |
| Db | 301 | DQPGSYSCMC | T | G | Y | R | L | A | O | D | H | R | C | E | D | V | D | C | I | L | B | P | S | P | C | P | Q | R | C | V | N | T | Q | G | S | F | E | C | H | C | Y | P | N | D | L | V | D | G | 360 | | | | | | |
| QY | 361 | ECVEPVDP | C | F | R | A | N | C | E | Y | O | C | O | P | L | N | O | T | S | Y | L | C | V | A | G | S | F | A | P | I | P | H | E | R | H | R | C | O | M | F | C | N | O | T | A | C | P | A | D | C | D | P | N | 420 | |
| Db | 361 | ECVEPVDP | C | F | R | A | N | C | E | Y | O | C | O | P | L | N | O | T | S | Y | L | C | V | A | G | S | F | A | P | I | P | H | E | R | H | R | C | O | M | F | C | N | O | T | A | C | P | A | D | C | D | P | N | 420 | |
| QY | 421 | TQASCECP | S | G | Y | I | L | D | G | F | I | C | T | D | I | D | E | C | E | N | G | G | F | C | S | G | V | C | H | N | L | P | G | T | F | E | C | I | C | G | P | D | S | A | L | A | R | H | I | G | T | D | C | 480 | |
| Db | 421 | TQASCECP | S | G | Y | I | L | D | G | F | I | C | T | D | I | D | E | C | E | N | G | G | F | C | S | G | V | C | H | N | L | P | G | T | F | E | C | I | C | G | P | D | S | A | L | A | R | H | I | G | T | D | C | 480 | |
| QY | 481 | DSGKVDG | D | G | S | G | S | G | E | P | P | S | T | P | T | G | S | T | L | T | P | P | A | V | G | L | V | H | S | G | 516 | | | | | | | | | | | | | | | | | | | | | | | | |
| Db | 481 | DSGKVDG | D | G | S | G | S | G | E | P | P | S | T | P | T | G | S | T | L | T | P | P | A | V | G | L | V | H | S | G | 516 | | | | | | | | | | | | | | | | | | | | | | | | |
| RESULT | | | 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Protein
LOCATION: 1..575
OTHER INFORMATION: /label= protein
OTHER INFORMATION: /note= "human thrombomodulin"
US-08-261-206A-59

Query Match 99.9%; Score 2912; DB 1; Length 575;
Best Local Similarity 99.8%; Pred. No. 1.3e-198;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGFPAEPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVGLALAGLGFPAEPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TYRSSVAADVISLLNGDGGVRRRLWIGLQPPGCGDKPRLGRLGFQWVTGDNNTSYS 120
DB 61 TYRSSVAADVISLLNGDGGVRRRLWIGLQPPGCGDKPRLGRLGFQWVTGDNNTSYS 120
QY 121 RWARLDNLGAPLCGPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180
DB 121 RWARLDNLGAPLCGPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180
QY 181 EPGAAAAVSITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAAVSITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
DB 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
QY 361 ECVEPVDPFCRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
DB 361 ECVEPVDPFCRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
QY 421 TQASCECEPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDALSARHIGTDC 480
DB 421 TQASCECEPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDALSARHIGTDC 480
QY 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 4

US-08-170-290A-54
Sequence 54, Application US/08170290A
Patent No. 5702931
GENERAL INFORMATION:
APPLICANT: Andrews, William H.
APPLICANT: Morser, Michael J.
APPLICANT: Ziehlender, Laura R.
TITLE OF INVENTION: No. 5702931el Mutagenesis Methods and
TITLE OF INVENTION: Compositions
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: James M. Heslin
STREET: 379 Lytton Ave.
CITY: Palo Alto
STATE: California

COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/170,290A
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05573
FILING DATE: 01-JUL-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/724,237
FILING DATE: 01-JUL-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 11972-58-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-170-290A-54

Query Match 99.7%; Score 2908; DB 1; Length 575;
Best Local Similarity 99.8%; Pred. No. 2.6e-198;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGFPAEPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVGLALAGLGFPAEPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TYRSSVAADVISLLNGDGGVRRRLWIGLQPPGCGDKPRLGRLGFQWVTGDNNTSYS 120
DB 61 TYRSSVAADVISLLNGDGGVRRRLWIGLQPPGCGDKPRLGRLGFQWVTGDNNTSYS 120
QY 121 RWARLDNLGAPLCGPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180
DB 121 RWARLDNLGAPLCGPLCAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAV 180
QY 181 EPGAAAAVSITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAAVSITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
DB 301 DQPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
QY 361 ECVEPVDPFCRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
DB 361 ECVEPVDPFCRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
QY 421 TQASCECEPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDALSARHIGTDC 480
DB 421 TQASCECEPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDALSARHIGTDC 480
QY 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516

Db 481 DSGKVDGSDSGEPPPTPGSTLTPPAVGLVHSG 516

RESULT 5

5256770-7
; Patent No. 5256770
; APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT,
; DAVID R.
; TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/506.325
; FILING DATE: 09-APR-1990
; SEQ ID NO: 7
; LENGTH: 572
5256770-7

Query Match 97.6%; Score 2846.5; DB 6; Length 572;
Best Local Similarity 98.6%; Pred. No. 5.7e-194;
Matches 509; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

QY 1 MGVLVIGALALAGFPAPAPBPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHLM 60
|||||
Db 1 MGVLVIGALALAGFPAPAPBPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHLM 60
|||||
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDPKRLGFGQWVTGDNNTSYS 120
|||||
Db 61 TVRSSVAADVISLLNGDGGV-RLWIGLQPLPGCGDPKRLGFGQWVTGDNNTSYS 119
|||||
QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 180
|||||
Db 120 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 179
|||||
QY 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPGAVQGHWAREAP 240
|||||
Db 180 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTA--GNVQGHWAREAP 237
|||||
QY 241 GAWDCSVENGGEHACNAPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
|||||
Db 238 GAWDCSVENGGEHACNAPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 297
|||||
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECYCPNYDLVDG 360
|||||
Db 298 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECYCPNYDLVDG 357
|||||
QY 361 ECVEPVPDFCFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCPN 420
|||||
Db 358 ECVEPVPDFCFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCPN 417
|||||
QY 421 TQASCEPGEYLLDDGFTCTDIDECENGFCGVCNHLPGTFECICGPDPSALARHIGTDC 480
|||||
Db 418 TQASCEPGEYLLDDGFTCTDIDECENGFCGVCNHLPGTFECICGPDPSALARHIGTDC 477
|||||
QY 481 DSGKVDGSDSGEPPPTPGSTLTPPAVGLVHSG 516
|||||
Db 478 DSGKVDGSDSGEPPPTPGSTLTPPAVGLVHSG 513
|||||

RESULT 6

US-08-733-564-2
; Sequence 2, Application US/08733564
; Patent No. 5916874
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, Kenji
; APPLICANT: MOCHIDA, Satoshi
; TITLE OF INVENTION: METHOD FOR TREATING LIVER INJURY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia

; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/733.564
; FILING DATE: 18 OCTOBER 1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 0216-0362P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 498 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-733-564-2

Query Match 96.9%; Score 2826; DB 2; Length 498;

Best Local Similarity 99.8%; Pred. No. 1.4e-192;

Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 APAEPQPGSGSQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 78
Db 1 APAEPQPGSGSQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 60
QY 79 GGVGRRRLWIGLQPLPGCGDPKRLGFGQWVTGDNNTSYSRWARLDLNGAPLCGPLCV 138
Db 61 GGVGRRRLWIGLQPLPGCGDPKRLGFGQWVTGDNNTSYSRWARLDLNGAPLCGPLCV 120
QY 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAVEPGAAAASVITYGTFFA 198
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAVEPGAAAASVITYGTFFA 180
QY 199 ARGADFQALPVGSSAAVAPLGLQLMCTAPGAVQGHWAREAPGAWDCSVENGGEHACNA 258
Db 181 ARGADFQALPVGSSAAVAPLGLQLMCTAPGAVQGHWAREAPGAWDCSVENGGEHACNA 240
QY 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMETGYRLAA 318
Db 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMETGYRLAA 300
QY 319 DQHRCEVDVDCILEPSPQRCVNTQGGFECYCPNYDLVDCGECYVPVDPDFCFRANCYQC 378
Db 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECYCPNYDLVDCGECYVPVDPDFCFRANCYQC 360
QY 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCPNTQASCEPGEYLLDDGFI 438
Db 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCPNTQASCEPGEYLLDDGFI 420
QY 439 CTDIDECENGFCGVCNHLPGTFECICGPDPSALARHIGTDCDSKVDGSDSGEPPPPS 498
Db 421 CTDIDECENGFCGVCNHLPGTFECICGPDPSALARHIGTDCDSKVDGSDSGEPPPPS 480
QY 499 PTPGSTLTPPAVGLVHSG 516
Db 481 PTPGSTLTPPAVGLVHSG 498
|||||

RESULT 7

US-08-312-870-3
; Sequence 3, Application US/08312870
; Patent No. 5639625

GENERAL INFORMATION:
APPLICANT: Carson, Craig W.
APPLICANT: Esmon, Charles T.
TITLE OF INVENTION: Method for Detecting Antibodies to
TITLE OF INVENTION: Thrombomodulin in Patients
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richards, Medlock & Andrews
STREET: 1201 Elm Street, Suite 4500
CITY: Dallas
STATE: Texas
COUNTRY: US
ZIP: 75270-2197
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/312,870
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hansen, Eugenia S.
REGISTRATION NUMBER: 31,966
REFERENCE/DOCKET NUMBER: OMRF B35150
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-939-4500
TELEFAX: 214-939-4600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-312-870-3

Query Match 96.8%; Score 2824; DB 1; Length 497;
Best Local Similarity 100.0%; Pred. No. 1.9e-192;
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 APAEPQGGSCVEHDFCFALYPCGATFLNASQICDGLRHLMTVRSVAADVLSLLNGD 78
|||||
DB 1 APAEPQGGSCVEHDFCFALYPCGATFLNASQICDGLRHLMTVRSVAADVLSLLNGD 60
|||||

QY 79 GGVGRRLTIGLQPPCGDPRKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGPLCV 138
|||||
DB 61 GGVGRRLTIGLQPPCGDPRKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGPLCV 120
|||||

QY 139 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAVEFGAAAAVSTYGTFFA 198
|||||
DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAVEFGAAAAVSTYGTFFA 180
|||||

QY 199 ARGADQALPVGSSAAVPLGLQMLCTAPPAGVQGHWAREAPAGWDCSVENGCEHACNA 258
|||||
DB 181 ARGADQALPVGSSAAVPLGLQMLCTAPPAGVQGHWAREAPAGWDCSVENGCEHACNA 240
|||||

QY 259 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQSGYSKMCETGYRLAA 318
|||||
DB 241 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQSGYSKMCETGYRLAA 300
|||||

QY 319 DQHRCEDDVDCILEPSPQPCQVNTGGFECYCNVLDLVGCEVPDPVDFCFRANCEYQC 378
|||||
DB 301 DQHRCEDDVDCILEPSPQPCQVNTGGFECYCNVLDLVGCEVPDPVDFCFRANCEYQC 360
|||||

QY 379 QPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQACPADCDPNTQASCECEPEGYLLDDGFI 438
|||||
DB 361 QPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQACPADCDPNTQASCECEPEGYLLDDGFI 420
|||||

QY 439 CTTDIDCENGGCGSVCHNLPGTFECICGPDSSALARIHIGTDCDSKGVGDGSGSGEPPPS 498
|||||
DB 421 CTTDIDCENGGCGSVCHNLPGTFECICGPDSSALARIHIGTDCDSKGVGDGSGSGEPPPS 480
|||||

QY 499 PTPGSTLTTPPAVGLVHS 515
|||||
DB 481 PTPGSTLTTPPAVGLVHS 497
|||||

RESULT 8
US-08-014-723-14
Sequence 14, Application US/08014723
Patent No. 5273962
GENERAL INFORMATION:
APPLICANT: Doi, Takeshi
APPLICANT: Iwasaki, Akio
APPLICANT: Saino, Yushi
APPLICANT: Kimura, Shigeru
APPLICANT: Ohkuchi, Masao
TITLE OF INVENTION: Thrombin-Binding Substance and Process
TITLE OF INVENTION: For Preparing the Same
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/014,723
FILING DATE: 19930208
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5273962man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-071-0 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)413-3000
TELEFAX: (703)413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-014-723-14

Query Match 94.9%; Score 2766; DB 1; Length 494;
Best Local Similarity 99.2%; Pred. No. 2.4e-188;
Matches 490; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVGLALALAGLGFPAEPQGGSCVEHDFCFALYPCGATFLNASQICDGLRHL 60
|||||
DB 1 MLGVLVGLALALAGLGFPAEPQGGSCVEHDFCFALYPCGATFLNASQICDGLRHL 60
|||||

QY 61 TVRSSVAADVLSLLNGDGGVGRRLTIGLQPPCGDPRKRLGRLGFQWVTGDNNTSY 120
|||||
DB 61 TVRSSVAADVLSLLNGDGGVGRRLTIGLQPPCGDPRKRLGRLGFQWVTGDNNTSY 120
|||||

QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 180
|||||
DB 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 180
|||||

QY 181 EPGAAAAVSTYGTTPFAARGADQALPVGSSAAVPLGLQMLCTAPPAGVQGHWAREAP 240
|||||
DB 181 EPGAAAAVSTYGTTPFAARGADQALPVGSSAAVPLGLQMLCTAPPAGVQGHWAREAP 240
|||||

QY 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
|||||

```
Db 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
Db 301 DQPGSYSCMETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCOPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
QY 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECIGPDSALARHIGTDC 480
Db 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECIGPDSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
Db 481 DSGKVEDYSGSGE 494

RESULT 9
US-08-110-011A-14
; Sequence 14, Application US/08110011A
; Patent No. 5354664
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,011A
; FILING DATE: 23-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5354664man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-110-011A-14

Query Match 94.9%; Score 2766; DB 1; Length 494;
Best Local Similarity 99.2%; Pred. No. 2.4e-188;
Matches 490; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVIGALAGLGFPAEPQGGSCQVCHDFALYPGPATFLNASQICDGLRGHLM 60
Db 1 MLGVLVIGALAGLGFPAEPQGGSCQVCHDFALYPGPATFLNASQICDGLRGHLM 60
```

```
QY 61 TYRSSVAADVISLLNGDGVGRRRLWIGLQPLPGGDKPKRLGRLGFQVMTGDNNTSYS 120
Db 61 TYRSSVAADVISLLNGDGVGRRRLWIGLQPLPGGDKPKRLGRLGFQVMTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGPICVAVSAEAATVPSEPIWEEQQCEVKADGFTCEHFFPATCRPLAV 180
Db 121 RWARLDLNGAPLCGPICVAVSAEAATVPSEPIWEEQQCEVKADGFTCEHFFPATCRPLAV 180
QY 181 EPGAAAAVITYGTPTFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVOGHWAREAP 240
Db 181 EPGAAAAVITYGTPTFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVOGHWAREAP 240
QY 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
Db 301 DQPGSYSCMETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCOPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
QY 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECIGPDSALARHIGTDC 480
Db 421 TOASCCEPBGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECIGPDSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
Db 481 DSGKVEDYSGSGE 494

RESULT 10
US-08-014-723-16
; Sequence 16, Application US/08014723
; Patent No. 5273962
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,723
; FILING DATE: 19930208
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5273962man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 494 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-014-723-16

Query Match          94.8%; Score 2764; DB 1; Length 494;
Best Local Similarity 99.0%; Pred. No. 3.3e-188;
Matches 489; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
   |||||
Db 1 MLGVLVLGALALAGLGFAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
   |||||

QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDPKRLGPRGFQWVTGDNNTSYS 120
   |||||
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDPKRLGPRGFQWVTGDNNTSYS 120
   |||||

QY 121 RWRDLNLCAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
   |||||
Db 121 RWRDLNLCAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
   |||||

QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP 240
   |||||
Db 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP 240
   |||||

QY 241 GAWDCSVENGCGCEHACNAIPGAPRCOCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
   |||||
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCOCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
   |||||

QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCPCQVNTQGGFECHCYPNYDLVDG 360
   |||||
Db 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCPCQVNTQGGFECHCYPNYDLVDG 360
   |||||

QY 361 ECVEPVDPFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
   |||||
Db 361 ECVEPVDPFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
   |||||

QY 421 TOASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTTEFCGPDSSALARHIGTDC 480
   |||||
Db 421 TOASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTTEFCGPDSSALARHIGTDC 480
   |||||

QY 481 DSGKYDGDGSGSGE 494
   |||||
Db 481 DSGKYDDEASGSGD 494
   |||||

RESULT 11
US-08-110-011A-16
; Sequence 16, Application US/08110011A
; Patent No. 5354564
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Okuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; City: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,011A
; FILING DATE: 23-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5354664man F.
; REGISTRATION NUMBER: 24, 618
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-110-011A-16

Query Match          94.8%; Score 2764; DB 1; Length 494;
Best Local Similarity 99.0%; Pred. No. 3.3e-188;
Matches 489; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
   |||||
Db 1 MLGVLVLGALALAGLGFAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
   |||||

QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDPKRLGPRGFQWVTGDNNTSYS 120
   |||||
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDPKRLGPRGFQWVTGDNNTSYS 120
   |||||

QY 121 RWRDLNLCAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
   |||||
Db 121 RWRDLNLCAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
   |||||

QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP 240
   |||||
Db 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP 240
   |||||

QY 241 GAWDCSVENGCGCEHACNAIPGAPRCOCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
   |||||
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCOCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
   |||||

QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCPCQVNTQGGFECHCYPNYDLVDG 360
   |||||
Db 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCPCQVNTQGGFECHCYPNYDLVDG 360
   |||||

QY 361 ECVEPVDPFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
   |||||
Db 361 ECVEPVDPFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
   |||||

QY 421 TOASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTTEFCGPDSSALARHIGTDC 480
   |||||
Db 421 TOASCECEGYILDDGFICTDIDECENGFCGSGVCHNLPGTTEFCGPDSSALARHIGTDC 480
   |||||

QY 481 DSGKYDGDGSGSGE 494
   |||||
Db 481 DSGKYDDEASGSGD 494
   |||||

RESULT 12
US-08-307-444A-2
; Sequence 2, Application US/08307444A
; Patent No. 5516659
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
; NUMBER OF SEQUENCES: 27
```

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307.444A
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,436
; FILING DATE: 26-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-307-444A-2

Query Match          92.2%; Score 2690; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 5.4e-183;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Db 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLPGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLPGFQWVTGDNNTSYS 120
Qy 121 RWARDLNGAPLCGPLCVAVAAEATVPSEPIWEBOQCEVKADGFLCFEFHPATCRPLAV 180
Db 121 RWARDLNGAPLCGPLCVAVAAEATVPSEPIWEBOQCEVKADGFLCFEFHPATCRPLAV 180
Qy 181 EPGAAAAVSYTGTFFAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAAVSYTGTFFAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVQGHWAREAP 240

Query Match          92.2%; Score 2690; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 5.4e-183;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Db 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLPGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLPGFQWVTGDNNTSYS 120
Qy 121 RWARDLNGAPLCGPLCVAVAAEATVPSEPIWEBOQCEVKADGFLCFEFHPATCRPLAV 180
Db 121 RWARDLNGAPLCGPLCVAVAAEATVPSEPIWEBOQCEVKADGFLCFEFHPATCRPLAV 180
Qy 181 EPGAAAAVSYTGTFFAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAAVSYTGTFFAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVQGHWAREAP 240

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307.444A
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/587.389
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/307.444
; FILING DATE: 19-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-587-389-2

Query Match          92.2%; Score 2690; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 5.4e-183;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Db 1  MLGVLVLGALALAGLGFPAEPQPGGQCVEHDCFCALYPGPATFLNASQICDGLRGHLM 60
Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLPGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPPKRLGRLPGFQWVTGDNNTSYS 120
Qy 121 RWARDLNGAPLCGPLCVAVAAEATVPSEPIWEBOQCEVKADGFLCFEFHPATCRPLAV 180
Db 121 RWARDLNGAPLCGPLCVAVAAEATVPSEPIWEBOQCEVKADGFLCFEFHPATCRPLAV 180
Qy 181 EPGAAAAVSYTGTFFAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAAVSYTGTFFAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVQGHWAREAP 240
```

QY 241 GAWDCSVNGGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVNGGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
Db 301 DQPGSYSCMETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVEPVDPFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
Db 361 ECVEPVDPFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
QY 421 TQASCECEGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECICGPDPSALARH 475
Db 421 TQASCECEGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECICGPDPSALARH 475
RESULT 14
US-08-307-444A-1
; Sequence 1, Application US/08307444A
; Patent No. 5516659
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,444A
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,436
; FILING DATE: 26-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-307-444A-1
Query Match 92.1%; Score 2686; DB 1; Length 475;
Best Local Similarity 99.8%; Pred. No. 1e-182;
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 MLGVLVLGALALAGLFPAPAEPPQGGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVLSLLNGDGGVGRRLWIGLQLPFGCDPKRLGRLGCFQVNTGDNNTSYS 120
Db 61 TVRSSVAADVLSLLNGDGGVGRRLWIGLQLPFGCDPKRLGRLGCFQVNTGDNNTSYS 120
QY 121 RWARLDNLGAPLCGPLCVAVSAAEAATVPSEPIWEEQOCEVKADGFLCEHFFPATCRPLAV 180
Db 121 RWARLDNLGAPLCGPLCVAVSAAEAATVPSEPIWEEQOCEVKADGFLCEHFFPATCRPLAV 180
QY 181 EPGAAAAYSVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
Db 181 EPGAAAAYSVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVNGGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVNGGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
Db 301 DQPGSYSCMETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVEPVDPFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
Db 361 ECVEPVDPFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
QY 421 TQASCECEGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECICGPDPSALARH 475
Db 421 TQASCECEGYILDDGFICTDIDECENGFGCSGVCHNLPGTTECICGPDPSALARH 475
RESULT 15
US-08-587-389-1
; Sequence 1, Application US/08587389
; Patent No. 5695964
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,389
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/307,444
; FILING DATE: 19-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-587-389-1

Query Match 92.1%; Score 2686; DB 1; Length 475;
Best Local Similarity 99.8%; Pred. No. 1e-182;
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| Qy | 1 | MLGVLVLAGALAGLGFAPAEPPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM | 60 |
| Db | 1 | MLGVLVLAGALAGLGFAPAEPPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM | 60 |
| Qy | 61 | TVRSSAADVISLLNGDGGVRRRLWTGLQPPGCGDPKRLGPLRGFQWVTGDNNTSYS | 120 |
| Db | 61 | TVRSSAADVISLLNGDGGVRRRLWTGLQPPGCGDPKRLGPLRGFQWVTGDNNTSYS | 120 |
| Qy | 121 | RWARDLNGAPLCGLCVAVSAAEATVPSEPIWEEQQCEVKADGFLCFEHPFATCRPLAV | 180 |
| Db | 121 | RWARDLNGAPLCGLCVAVSAAEATVPSEPIWEEQQCEVKADGFLCFEHPFATCRPLAV | 180 |
| Qy | 181 | EPGAAAASITYGTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVOGHWAREAP | 240 |
| Db | 181 | EPGAAAASITYGTPFAARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVOGHWAREAP | 240 |
| Qy | 241 | GAWDCSVENGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP | 300 |
| Db | 241 | GAWDCSVENGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP | 300 |
| Qy | 301 | DOPGSYSCMCEYQCPNLQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN | 360 |
| Db | 301 | DOPGSYSCMCEYQCPNLQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN | 360 |
| Qy | 361 | ECVEPVDFCFRANCEYQCPNLQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN | 420 |
| Db | 361 | ECVEPVDFCFRANCEYQCPNLQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN | 420 |
| Qy | 421 | TQASCEPEGYILDDGFICTDIDECENGCGCGVCHNLPGTFECICGPDSDALAH | 475 |
| Db | 421 | TQASCEPEGYILDDGFICTDIDECENGCGCGVCHNLPGTFECICGPDSDALVRH | 475 |

Search completed: May 8, 2002, 12:39:21
Job time: 212 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:35:49 ; Search time 21.32 Seconds
(without alignments)
544.639 Million cell updates/sec

Title: US-09-509-994-1
Perfect score: 2916
Sequence: 1 MGVVLGALAGLGPAP.....PSPTPGSLTPPAVGLVHSG 516

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents:AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|---------------------|--------------------|
| 1 | 2916 | 100.0 | 575 | 1 US-08-261-206A-59 | Sequence 59, Appl |
| 2 | 2912 | 99.9 | 575 | 1 US-08-312-870-1 | Sequence 1, Appl |
| 3 | 2912 | 99.9 | 575 | 6 5466668-6 | Patent No. 5466668 |
| 4 | 2904 | 99.6 | 575 | 1 US-08-170-290A-54 | Sequence 54, Appl |
| 5 | 2842.5 | 97.5 | 572 | 6 5256770-7 | Patent No. 5256770 |
| 6 | 2830 | 97.1 | 498 | 2 US-08-733-564-2 | Sequence 2, Appl |
| 7 | 2830 | 96.7 | 497 | 1 US-08-312-870-3 | Sequence 3, Appl |
| 8 | 2770 | 95.0 | 494 | 1 US-08-014-723-14 | Sequence 14, Appl |
| 9 | 2770 | 95.0 | 494 | 1 US-08-110-011A-14 | Sequence 14, Appl |
| 10 | 2768 | 94.9 | 494 | 1 US-08-014-723-16 | Sequence 16, Appl |
| 11 | 2768 | 94.9 | 494 | 1 US-08-110-011A-16 | Sequence 16, Appl |
| 12 | 2690 | 92.2 | 475 | 1 US-08-307-444A-1 | Sequence 1, Appl |
| 13 | 2690 | 92.2 | 475 | 1 US-08-587-389-1 | Sequence 1, Appl |
| 14 | 2686 | 92.1 | 475 | 1 US-08-307-444A-2 | Sequence 2, Appl |
| 15 | 2686 | 92.1 | 475 | 1 US-08-587-389-2 | Sequence 2, Appl |
| 16 | 2684 | 92.0 | 476 | 1 US-08-014-723-1 | Sequence 1, Appl |
| 17 | 2684 | 92.0 | 476 | 1 US-08-110-011A-1 | Sequence 1, Appl |
| 18 | 2682 | 92.0 | 476 | 1 US-08-014-723-2 | Sequence 2, Appl |
| 19 | 2682 | 92.0 | 476 | 1 US-08-014-723-18 | Sequence 18, Appl |
| 20 | 2682 | 92.0 | 476 | 1 US-08-110-011A-2 | Sequence 2, Appl |
| 21 | 2682 | 92.0 | 476 | 1 US-08-110-011A-18 | Sequence 18, Appl |
| 22 | 2596 | 89.0 | 456 | 1 US-08-307-444A-3 | Sequence 3, Appl |
| 23 | 2596 | 89.0 | 456 | 1 US-08-587-389-3 | Sequence 3, Appl |
| 24 | 2592 | 88.9 | 456 | 1 US-08-307-444A-4 | Sequence 4, Appl |
| 25 | 2592 | 88.9 | 456 | 1 US-08-587-389-4 | Sequence 4, Appl |
| 26 | 2543 | 87.2 | 446 | 1 US-08-307-444A-5 | Sequence 5, Appl |
| 27 | 2543 | 87.2 | 446 | 1 US-08-587-389-5 | Sequence 5, Appl |

| | | | | | | |
|----|-------|------|------|---|-------------------|--------------------|
| 28 | 1617 | 55.5 | 275 | 1 | US-08-312-870-7 | Sequence 7, Appl |
| 29 | 1159 | 39.7 | 215 | 1 | US-08-312-870-5 | Sequence 5, Appl |
| 30 | 685 | 23.5 | 114 | 2 | US-08-733-564-1 | Sequence 1, Appl |
| 31 | 685 | 23.5 | 115 | 1 | US-08-312-870-9 | Sequence 9, Appl |
| 32 | 580 | 19.9 | 652 | 2 | US-08-751-305-2 | Sequence 2, Appl |
| 33 | 358 | 12.3 | 638 | 2 | US-08-897-443-1 | Sequence 1, Appl |
| 34 | 352 | 12.1 | 58 | 1 | US-08-261-206A-3 | Sequence 3, Appl |
| 35 | 328.5 | 11.3 | 956 | 2 | US-08-897-443-3 | Sequence 3, Appl |
| 36 | 323 | 11.1 | 1253 | 3 | US-08-479-722B-4 | Sequence 4, Appl |
| 37 | 320 | 11.0 | 1394 | 6 | 5177197-30 | Patent No. 5177197 |
| 38 | 319.5 | 11.0 | 1833 | 3 | US-08-479-722B-2 | Sequence 2, Appl |
| 39 | 319.5 | 11.0 | 1833 | 5 | PCT-US95-02251-18 | Sequence 18, Appl |
| 40 | 315.5 | 10.8 | 443 | 2 | US-08-833-963C-2 | Sequence 2, Appl |
| 41 | 315.5 | 10.8 | 443 | 3 | US-08-980-514-1 | Sequence 1, Appl |
| 42 | 308 | 10.6 | 448 | 2 | US-08-884-072-1 | Sequence 1, Appl |
| 43 | 308 | 10.6 | 448 | 4 | US-09-212-168-1 | Sequence 1, Appl |
| 44 | 303.5 | 10.4 | 1251 | 5 | PCT-US95-02251-3 | Sequence 3, Appl |
| 45 | 303.5 | 10.4 | 1252 | 1 | US-08-199-780-3 | Sequence 3, Appl |

ALIGNMENTS

RESULT 1
US-08-261-206A-59
; Sequence 59, Application US/08261206A
; Patent No. 5574007
; GENERAL INFORMATION:
; APPLICANT: Zushi, Mitichitaka
; APPLICANT: Gomi, Komakazu
; APPLICANT: Yamamoto, Shuji
; APPLICANT: Suzuki, Koji
; APPLICANT: Matsuda, Akio
; TITLE OF INVENTION: A Polypeptide Capable of Interacting
; TITLE OF INVENTION: with Thrombin
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261.206A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/740,492
; FILING DATE: 03-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30330
; REFERENCE/DOCKET NUMBER: 216-275P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: Protein
; LOCATION: 1..575
; OTHER INFORMATION: /label= protein
; OTHER INFORMATION: /note= "human thrombomodulin"
US-08-261-206A-59

Query Match 100.0%; Score 2916; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 1e-198;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVLVLGALALAGLGFAPAPPEPGGSCQVEHDFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MGVLVLGALALAGLGFAPAPPEPGGSCQVEHDFALYPGPATFLNASQICDGLRGHLM 60

QY 61 TVRSSVAADVLSLLNGDGGVRRRLWTGLQLPPCGDPKRLGPLRGFWQVWTDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWTGLQLPPCGDPKRLGPLRGFWQVWTDNNTSYS 120

QY 121 RWARDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQQCEVKADGFLCEFHFPATCRPLAV 180
DB 121 RWARDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQQCEVKADGFLCEFHFPATCRPLAV 180

QY 181 EPGAAAAVSTYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240
DB 181 EPGAAAAVSTYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240

QY 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300

QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360
DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360

QY 361 ECVEPVPDFCFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420
DB 361 ECVEPVPDFCFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420

QY 421 TOACSECEPGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALSALVRHIGTDC 480
DB 421 TOACSECEPGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALSALVRHIGTDC 480

QY 481 DSGKVDGSDSGSGEPPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGEPPPTPGSTLTTPPAVGLVHSG 516

RESULT 2
US-08-312-870-1
; Sequence 1, Application US/08312870
; Patent No. 5639625
; GENERAL INFORMATION:
; APPLICANT: Carson, Craig W.
; APPLICANT: Esmon, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to
; TITLE OF INVENTION: Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Richards, Medlock & Andrews
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: Texas
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,870
; FILING DATE:

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hansen, Eugenia S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: OMRF B35150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 19..575
US-08-312-870-1

Query Match 99.9%; Score 2912; DB 1; Length 575;
Best Local Similarity 99.8%; Pred. No. 1.9e-198;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLVLGALALAGLGFAPAPPEPGGSCQVEHDFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MGVLVLGALALAGLGFAPAPPEPGGSCQVEHDFALYPGPATFLNASQICDGLRGHLM 60

QY 61 TVRSSVAADVLSLLNGDGGVRRRLWTGLQLPPCGDPKRLGPLRGFWQVWTDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWTGLQLPPCGDPKRLGPLRGFWQVWTDNNTSYS 120

QY 121 RWARDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQQCEVKADGFLCEFHFPATCRPLAV 180
DB 121 RWARDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQQCEVKADGFLCEFHFPATCRPLAV 180

QY 181 EPGAAAAVSTYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240
DB 181 EPGAAAAVSTYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240

QY 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300

QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360
DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360

QY 361 ECVEPVPDFCFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420
DB 361 ECVEPVPDFCFRANCYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420

QY 421 TOACSECEPGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALSALVRHIGTDC 480
DB 421 TOACSECEPGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALSALVRHIGTDC 480

QY 481 DSGKVDGSDSGSGEPPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGEPPPTPGSTLTTPPAVGLVHSG 516

RESULT 3
5466668-6
; Patent No. 5466668
; APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT,
; DAVID R.
; TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR
; PHARMACEUTICAL USE
; NUMBER OF SEQUENCES: 57
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/155,346
FILING DATE: 22-NOV-1993
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION NUMBER: 568,456
FILING DATE: 15-AUG-1990
APPLICATION NUMBER: 506,325
FILING DATE: 09-APR-1990
APPLICATION NUMBER: 406,941
FILING DATE: 13-SEP-1989
APPLICATION NUMBER: 345,374
FILING DATE: 28-APR-1989
SEQ ID NO: 6
LENGTH: 575
5466668-6

Query Match 99.9%; Score 2912; DB 6; Length 575;
Best Local Similarity 99.8%; Pred. No. 1.9e-198;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLVLGALALAGLGFPAEPQPGGSCQVHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MGVLVLGALALAGLGFPAEPQPGGSCQVHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPRKRLGRLGFQVWTDGNNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPRKRLGRLGFQVWTDGNNNTSYS 120
QY 121 RWRDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFPATCRPLAV 180
DB 121 RWRDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFPATCRPLAV 180
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
QY 361 ECVEPVDPFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
DB 361 ECVEPVDPFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
QY 421 TQASCECPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALVRHIGTDC 480
DB 421 TQASCECPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 4
US-08-170-290A-54
Sequence 54, Application US/08170290A
Patent No. 5702931
GENERAL INFORMATION:
APPLICANT: Andrews, William H.
APPLICANT: Morser, Michael J.
APPLICANT: Ziehlender, Laura R.
TITLE OF INVENTION: No. 5702931el Mutagenesis Methods and
TITLE OF INVENTION: Compositions
NUMBER OF INVENTIONS: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: James M. Heslin
STREET: 379 Lytton Ave.
CITY: Palo Alto
STATE: California

COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/170,290A
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05573
FILING DATE: 01-JUL-1992
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/724,237
FILING DATE: 01-JUL-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 11972-58-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-170-290A-54

Query Match 99.6%; Score 2904; DB 1; Length 575;
Best Local Similarity 99.6%; Pred. No. 7.1e-198;
Matches 514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MGVLVLGALALAGLGFPAEPQPGGSCQVHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MGVLVLGALALAGLGFPAEPQPGGSCQVHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPRKRLGRLGFQVWTDGNNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPCGDPRKRLGRLGFQVWTDGNNNTSYS 120
QY 121 RWRDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFPATCRPLAV 180
DB 121 RWRDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFPATCRPLAV 180
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
QY 361 ECVEPVDPFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
DB 361 ECVEPVDPFCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
QY 421 TQASCECPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALVRHIGTDC 480
DB 421 TQASCECPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516
|||||

Db 481 DSGKVDGDSGSGEPPSPPTPGSTLTTPAVGLVHSG 516

RESULT 5
5256770-7
; Patent No. 5256770
; APPLICANT: GLASER, CHARLES B.; MORSER, MICHAEL J.; LIGHT,
; DAVID R.
; TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/506,325
; FILING DATE: 09-APR-1990
; SEQ ID NO: 7:
; LENGTH: 572
5256770-7

Query Match 97.58; Score 2842.5; DB 6; Length 572;
Best Local Similarity 98.4%; Pred. No. 1.6e-193;
Matches 508; Conservative 0; Mismatches 5; Indels 3; Gaps 2;
QY 1 MGVLVIGALALAGLPPAPAEPOGSGOCVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MGVLVIGALALAGLPPAPAEPOGSGOCVEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWTGLPPGCGDKPKRLGRLGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGV--RRLWTGLPPGCGDKPKRLGRLGFQWVTGDNNTSYS 119
QY 121 RWARLDLNGAPLCPCLVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 180
DB 120 RWARLDLNGAPLCPCLVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 179
QY 181 EPGAAAASVTYGTPTFAARGADQFALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 180 EPGAAAASVTYGTPTFAARGADQFALPVGSSAAVAPLGLQLMCTA--GNVQGHWAREAP 237
QY 241 GAWDCSVENGSGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPNP 300
DB 238 GAWDCSVENGSGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPNP 297
QY 301 DQPGSYSCMCTGYRLAADQHRCEBDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
DB 298 DQPGSYSCMCTGYRLAADQHRCEBDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 357
QY 361 ECVEPVDPCEFRANCEYQCPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCPN 420
DB 358 ECVEPVDPCEFRANCEYQCPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCPN 417
QY 421 TQASCEGPEGYLLDDGFICTIDDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDC 480
DB 418 TQASCEGPEGYLLDDGFICTIDDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDC 477
QY 481 DSGKVDGDSGSGEPPSPPTPGSTLTTPAVGLVHSG 516
DB 478 DSGKVDGDSGSGEPPSPPTPGSTLTTPAVGLVHSG 513

RESULT 6
US-08-733-564-2
; Sequence 2, Application us/08733564
; Patent No. 5916874
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, Kenji
; APPLICANT: MOCHIDA, Satoshi
; TITLE OF INVENTION: METHOD FOR TREATING LIVER INJURY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia

; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/733,564
; FILING DATE: 18 OCTOBER 1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 0216-0362P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 498 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-733-564-2

Query Match 97.18; Score 2830; DB 2; Length 498;

Best Local Similarity 100.0%; Pred. No. 1e-192;

Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 APAPQPQGGSCVCEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVISLLNGD 78
DB 1 APAPQPQGGSCVCEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVISLLNGD 60
QY 79 GGVGRRRLWTGLQPLPPGCGDKPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCPCLV 138
DB 61 GGVGRRRLWTGLQPLPPGCGDKPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCPCLV 120
QY 139 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAVEPGAAAAVSIYGTGPPA 198
DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAVEPGAAAAVSIYGTGPPA 180
QY 199 ARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGSGEHCACNA 258
DB 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGSGEHCACNA 240
QY 259 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPNPDPGYSYSCMCTGYRLAA 318
DB 241 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPNPDPGYSYSCMCTGYRLAA 300
QY 319 DQHRCEBDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDGECVPEVDPCEFRANCEYQC 378
DB 301 DQHRCEBDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDGECVPEVDPCEFRANCEYQC 360
QY 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCPNTQASCEGPEGYLLDDGFI 438
DB 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCPNTQASCEGPEGYLLDDGFI 420
QY 439 CTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDCDCKVDDGSGDGGSGEPPPPS 498
DB 421 CTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDCDCKVDDGSGDGGSGEPPPPS 480
QY 499 PTPGSTLTTPAVGLVHSG 516
DB 481 PTPGSTLTTPAVGLVHSG 498

RESULT 7
US-08-312-870-3
; Sequence 3, Application US/08312870
; Patent No. 5639625

GENERAL INFORMATION:
APPLICANT: Carson, Craig W.
APPLICANT: Esmon, Charles T.
TITLE OF INVENTION: Method for Detecting Antibodies to
TITLE OF INVENTION: Thrombomodulin in Patients
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richards, Medlock & Andrews
STREET: 1201 Elm Street, Suite 4500
CITY: Dallas
STATE: Texas
COUNTRY: US
ZIP: 75270-2197
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/312,870
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hansen, Eugenia S.
REGISTRATION NUMBER: 31,966
REFERENCE/DOCKET NUMBER: OMRF B35150
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-939-4500
TELEFAX: 214-939-4600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-312-870-3

Query Match 96.7%; Score 2820; DB 1; Length 497;
Best Local Similarity 99.8%; Pred. No. 5.2e-192;
Matches 496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRLMTVRSVAADVLSLLNGD 78
|||||
DB 1 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRLMTVRSVAADVLSLLNGD 60
|||||
QY 79 GGVRRRLWIGLQPLPGCGDPRKRLGFRQWVTGDNNTSYSRWALDLNGAPLCGLCV 138
|||||
DB 61 GGVRRRLWIGLQPLPGCGDPRKRLGFRQWVTGDNNTSYSRWALDLNGAPLCGLCV 120
|||||
QY 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPAGAAAASITYGTPFA 198
|||||
DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPAGAAAASITYGTPFA 180
|||||
QY 199 ARGADFQALPVGSSAAVAPLGLQMLCTAPGAVQGHWAREAGAWDCSVENGCGEACNA 258
|||||
DB 181 ARGADFQALPVGSSAAVAPLGLQMLCTAPGAVQGHWAREAGAWDCSVENGCGEACNA 240
|||||
QY 259 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPSYSCMCTGYRLAA 318
|||||
DB 241 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPSYSCMCTGYRLAA 300
|||||
QY 319 DQHRCEVDVDCILEPSPCPQRCVNTGGFECYCNVLDVDCVBPVDPQCFRANCEYQC 378
|||||
DB 301 DQHRCEVDVDCILEPSPCPQRCVNTGGFECYCNVLDVDCVBPVDPQCFRANCEYQC 360
|||||
QY 379 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPNTQASCECEGYILDDGFI 438
|||||
DB 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPNTQASCECEGYILDDGFI 420
|||||
QY 439 CTTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDSGKVGDSGSGGPPPS 498
|||||
DB 421 CTTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDSGKVGDSGSGGPPPS 480
|||||

QY 499 PTPGSTLTTPPAVGLVHS 515
|||||
DB 481 PTPGSTLTTPPAVGLVHS 497
|||||
RESULT 8
US-08-014-723-14
Sequence 14, Application US/08014723
Patent No. 5273962
GENERAL INFORMATION:
APPLICANT: Doi, Takeshi
APPLICANT: Iwasaki, Akio
APPLICANT: Saino, Yushi
APPLICANT: Kimura, Shigeru
APPLICANT: Ohkuchi, Masao
TITLE OF INVENTION: Thrombin-Binding Substance and Process
TITLE OF INVENTION: For Preparing the Same
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/014,723
FILING DATE: 19930208
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5273962man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-071-0 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)413-3000
TELEFAX: (703)413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-014-723-14

Query Match 95.0%; Score 2770; DB 1; Length 494;
Best Local Similarity 99.4%; Pred. No. 1.8e-188;
Matches 491; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLPAPAEPOPGSCQVEHDCFALYGPATFLNASQICDGLRHLM 60
|||||
DB 1 MLGVLVLGALALAGLPAPAEPOPGSCQVEHDCFALYGPATFLNASQICDGLRHLM 60
|||||
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPRKRLGFRQWVTGDNNTSYS 120
|||||
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPRKRLGFRQWVTGDNNTSYS 120
|||||
QY 121 RWRDLNGLAPLCGLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180
|||||
DB 121 RWRDLNGLAPLCGLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180
|||||
QY 181 EPGAAAASVITYGTPEAARGADFOALPVGSSAAVAPLGLQMLCTAPGAVQGHWAREAP 240
|||||
DB 181 EPGAAAASVITYGTPEAARGADFOALPVGSSAAVAPLGLQMLCTAPGAVQGHWAREAP 240
|||||
QY 241 GAWDCSVENGCGEACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
|||||

```
Db 241 GAWDCSVENGCEHACNATPGAPRCQCPAGAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
QY 421 TOASCCEPBGYILDDGFICTDIDECENGFCGVCVCHNLPGTTECICGPDPSALVRHIGTDC 480
Db 421 TOASCCEPBGYILDDGFICTDIDECENGFCGVCVCHNLPGTTECICGPDPSALVRHIGTDC 480
QY 481 DSGKVDGDDGSGGE 494
Db 481 DSGKVEDYSGSGE 494

RESULT 9
US-08-110-011A-14
; Sequence 14, Application US/08110011A
; Patent No. 5354664
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110.011A
; FILING DATE: 23-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5354664man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-110-011A-14

Query Match 95.0%; Score 2770; DB 1; Length 494;
Best Local Similarity 99.4%; Pred. No. 1.8e-188;
Matches 491; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVVLVGLALAGLGFAPAPQPGSQVCVCHNLPGTTECICGPDPSALVRHIGTDCGLRHLM 60
Db 1 MGVVLVGLALAGLGFAPAPQPGSQVCVCHNLPGTTECICGPDPSALVRHIGTDCGLRHLM 60
```

```
QY 61 TVRSSVAADVISLLNGDGVGRRRLWIGLQPLPGCGDPRKLGRLGFGOWVTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGVGRRRLWIGLQPLPGCGDPRKLGRLGFGOWVTGDNNTSYS 120
QY 121 RWARDLNGAPLCGPLCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180
Db 121 RWARDLNGAPLCGPLCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180
QY 181 EFGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVOGHWAREAP 240
Db 181 EFGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVOGHWAREAP 240
QY 241 GAWDCSVENGCEHACNATPGAPRCQCPAGAALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCEHACNATPGAPRCQCPAGAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
QY 421 TOASCCEPBGYILDDGFICTDIDECENGFCGVCVCHNLPGTTECICGPDPSALVRHIGTDC 480
Db 421 TOASCCEPBGYILDDGFICTDIDECENGFCGVCVCHNLPGTTECICGPDPSALVRHIGTDC 480
QY 481 DSGKVDGDDGSGGE 494
Db 481 DSGKVEDYSGSGE 494

RESULT 10
US-08-014-723-16
; Sequence 16, Application US/08014723
; Patent No. 5273962
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014.723
; FILING DATE: 19930208
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5273962man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 494 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-014-723-16

Query Match 94.9%; Score 2768; DB 1; Length 494;
Best Local Similarity 99.2%; Pred. No. 2.4e-188;
Matches 490; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFPPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60
DB 1 MLGVLVLGALALAGLGFPPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180
DB 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180
QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQGSYSVMCTGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECCHYCNIDLVGD 360
DB 301 DQGSYSVMCTGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECCHYCNIDLVGD 360
QY 361 ECVEPVDPCFRANCEYOCPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
DB 361 ECVEPVDPCFRANCEYOCPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
QY 421 TQASCEPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
DB 421 TQASCEPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
DB 481 DSGKVDDEASGSD 494

RESULT 11

US-08-110-011A-16
Sequence 16, Application US/08110011A
Patent No. 5354664

GENERAL INFORMATION:

APPLICANT: Doi, Takeshi
APPLICANT: Iwasaki, Akio
APPLICANT: Saino, Yushi
APPLICANT: Kimura, Shigeru
APPLICANT: Okuchi, Masao
TITLE OF INVENTION: Thrombin-Binding Substance and Process
TITLE OF INVENTION: For Preparing the Same
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/110.011A
FILING DATE: 23-AUG-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5354664man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-073-0 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)413-3000
TELEFAX: (703)413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-110-011A-16

Query Match 94.9%; Score 2768; DB 1; Length 494;
Best Local Similarity 99.2%; Pred. No. 2.4e-188;
Matches 490; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFPPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60
DB 1 MLGVLVLGALALAGLGFPPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180
DB 121 RWARLDLNGAPLCGLPCLVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAV 180
QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQGSYSVMCTGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECCHYCNIDLVGD 360
DB 301 DQGSYSVMCTGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECCHYCNIDLVGD 360
QY 361 ECVEPVDPCFRANCEYOCPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
DB 361 ECVEPVDPCFRANCEYOCPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPN 420
QY 421 TQASCEPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
DB 421 TQASCEPEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
DB 481 DSGKVDDEASGSD 494

RESULT 12

US-08-307-444A-1
Sequence 1, Application US/08307444A
Patent No. 5516659

GENERAL INFORMATION:

APPLICANT: NII, ATSUSHI
APPLICANT: MORISHITA, HIDEAKI
APPLICANT: UEMURA, AKIO
APPLICANT: MOCHIDA, EI
TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
NUMBER OF SEQUENCES: 27

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,444A
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,436
; FILING DATE: 26-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-307-444A-1

Query Match          92.2%; Score 2690; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 7.6e-183;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVGLALALAGLGFAPAEPPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVGLALALAGLGFAPAEPPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVVISLLNGDGGVGRRLRWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVVISLLNGDGGVGRRLRWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180
DB 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180
QY 181 EPGAAAAAIVSYTGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240
DB 181 EPGAAAAAIVSYTGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHCACNALPGAPRCOCAGALQADGRSCTASATOSCDNLCEHFCVNP 300
DB 241 GAWDCSVENGCGEHCACNALPGAPRCOCAGALQADGRSCTASATOSCDNLCEHFCVNP 300
QY 301 DQPGSYSCMCETGYRLAAQDHRCEVDYDCCILEPSPCQRCVNTQGGFECHCYPNYDLVDG 360
DB 301 DQPGSYSCMCETGYRLAAQDHRCEVDYDCCILEPSPCQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVEPVDPFRANCEYQCQPLNQTSLYLCVCAEGFADIPHEPHRCQMFQNTACPADCDN 420
DB 361 ECVEPVDPFRANCEYQCQPLNQTSLYLCVCAEGFADIPHEPHRCQMFQNTACPADCDN 420
QY 421 TQASCECPGEGYILDDGFICTDDIDECENGCGFCGCHNLPGTFCICGPDPSALVRH 475
DB 421 TQASCECPGEGYILDDGFICTDDIDECENGCGFCGCHNLPGTFCICGPDPSALVRH 475

RESULT 13
US-08-587-389-1
; Sequence 1, Application US/08587389
; Patent No. 5695964
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT
; TITLE OF INVENTION: PRODUCTION THEREOF, AND THERAPEUTIC AGENT (AS AMENDED)
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,389
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/307,444
; FILING DATE: 19-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-587-389-1

Query Match          92.2%; Score 2690; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 7.6e-183;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVGLALALAGLGFAPAEPPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVGLALALAGLGFAPAEPPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVVISLLNGDGGVGRRLRWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVVISLLNGDGGVGRRLRWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120
QY 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180
DB 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180
QY 181 EPGAAAAAIVSYTGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240
DB 181 EPGAAAAAIVSYTGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240
```

```
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
|||||
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECCHCYPNYDLVDG 360
|||||
Db 301 DQPGSYSCMETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECCHCYPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
|||||
Db 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
QY 421 TOASCCEPGYILDDGFICTDIDECENGCGFCGVCCHNLPGTFECICGPDALVRH 475
|||||
Db 421 TOASCCEPGYILDDGFICTDIDECENGCGFCGVCCHNLPGTFECICGPDALVRH 475

RESULT 14
US-08-307-444A-2
; Sequence 2, Application US/08307444A
; Patent No. 5516659
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,444A
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,436
; FILING DATE: 26-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-307-444A-2

Query Match 92.1%; Score 2686; DB 1; Length 475;
Best Local Similarity 99.8%; Pred. No. 1.5e-182;
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLGVLVGLALAGLGFPAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60
```

```
|||||
Db 1 MLGVLVGLALAGLGFPAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60
QY 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCCDPRKRLGRLGFQWVTTGNNTSYS 120
|||||
Db 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCCDPRKRLGRLGFQWVTTGNNTSYS 120
QY 121 RWARLDLNGAPLCGPLCVASAAEAATVPSEPIWEEQOCEVKADGFLCEFFHPATCRPLAV 180
|||||
Db 121 RWARLDLNGAPLCGPLCVASAAEAATVPSEPIWEEQOCEVKADGFLCEFFHPATCRPLAV 180
QY 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAGVQGHWAREAP 240
|||||
Db 181 EPGAAAASVITYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAGVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
|||||
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECCHCYPNYDLVDG 360
|||||
Db 301 DQPGSYSCMETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECCHCYPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
|||||
Db 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
QY 421 TOASCCEPGYILDDGFICTDIDECENGCGFCGVCCHNLPGTFECICGPDALVRH 475
|||||
Db 421 TOASCCEPGYILDDGFICTDIDECENGCGFCGVCCHNLPGTFECICGPDALVRH 475

RESULT 15
US-08-587-389-2
; Sequence 2, Application US/08587389
; Patent No. 5695964
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,389
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/307,444
; FILING DATE: 19-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A.
; REGISTRATION NUMBER: 27,075
; REFERENCE/DOCKET NUMBER: JAO 27706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6400
; TELEFAX: (703) 836-2787
; TELEX: 90-1799 PTO ALEX
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```


; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-587-389-2

```
Query Match          92.1%; Score 2686; DB 1; Length 475;
Best Local Similarity 99.8%; Pred. No. 1.5e-182;
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLGVLVLGALALAGLGFAPAPQPGGQCVHEHDCFALYPGPATFELNASQICDGLRGHLM 60
Db 1 MLGVLVLGALALAGLGFAPAPQPGGQCVHEHDCFALYPGPATFELNASQICDGLRGHLM 60

Qy 61 TVRSSVAADVISLLNGDGGVGRRLWIGLQLPPCGDPPKRLGPLRGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVISLLNGDGGVGRRLWIGLQLPPCGDPPKRLGPLRGFQWVTGDNNTSYS 120

Qy 121 RWRDLNLGAPLCGLPCVAVSAAEATVPSEPIWEQQQCEVKADGFLCEFFHPATCRPLAV 180
Db 121 RWRDLNLGAPLCGLPCVAVSAAEATVPSEPIWEQQQCEVKADGFLCEFFHPATCRPLAV 180

Qy 181 EPGAAAAVSTIYTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
Db 181 EPGAAAAVSTIYTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240

Qy 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPNP 300
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPNP 300

Qy 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILLEPSPQRCVNTQGGFECCHYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILLEPSPQRCVNTQGGFECCHYPNYDLVDG 360

Qy 361 ECVPEVDFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420
Db 361 ECVPEVDFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420

Qy 421 TQASCEPGEYILDDGFICTDIDECENGFGCGVCHNLPGTFCICGPDLSALVRH 475
Db 421 TQASCEPGEYILDDGFICTDIDECENGFGCGVCHNLPGTFCICGPDLSALVRH 475
```

Search completed: May 8, 2002, 12:39:20
Job time: 211 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:35:19 ; Search time 42.86 Seconds
(without alignments)
891.783 Million cell updates/sec

Title: US-09-509-994-1
Perfect score: 2916
Sequence: 1 MLGVLVLGALAGLGFAP.....PSPTPGSTLTPPAVLGVHSG 516

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_1101.*

```

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result NO. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 2916 | 100.0 | 516 | 13 AAR22016 | Truncated human th |
| 2 | 2916 | 100.0 | 516 | 20 AAY09347 | Human thrombomodul |
| 3 | 2916 | 100.0 | 516 | 21 AAY83934 | Human thrombomodul |
| 4 | 2916 | 100.0 | 516 | 21 AAY69529 | Human thrombomodul |
| 5 | 2916 | 100.0 | 575 | 12 AAR11534 | Human thrombomodul |
| 6 | 2916 | 100.0 | 575 | 14 AAR41806 | Thrombomodulin. H |
| 7 | 2916 | 100.0 | 575 | 14 AAR43031 | Human thrombomodul |
| 8 | 2912 | 99.9 | 516 | 13 AAR22018 | Human thrombomodul |
| 9 | 2912 | 99.9 | 516 | 20 AAY09348 | Human thrombomodul |
| 10 | 2912 | 99.9 | 516 | 21 AAY83935 | Human thrombomodul |
| 11 | 2908 | 99.7 | 516 | 13 AAR22017 | Human thrombomodul |

| ID | Score | Query Match | Length | ID | Description |
|----|--------|-------------|--------|-------------|--------------------|
| 12 | 2906 | 99.7 | 515 | 12 AAR10617 | Soluble thrombomod |
| 13 | 2906 | 99.7 | 575 | 9 AAR22070 | Human thrombomodul |
| 14 | 2904 | 99.6 | 575 | 14 AAR31572 | Human thrombomodul |
| 15 | 2899 | 99.4 | 575 | 13 AAR20639 | Human urinary thro |
| 16 | 2875 | 98.6 | 575 | 20 AAR73970 | Human thrombomodul |
| 17 | 2858 | 98.0 | 575 | 13 AAR22189 | Sequence of thromb |
| 18 | 2830 | 97.1 | 498 | 16 AAR84185 | Human derived thro |
| 19 | 2830 | 97.1 | 498 | 18 AAR01600 | Thrombomodulin TME |
| 20 | 2825 | 96.9 | 498 | 21 AAY67402 | Novel sugar chain- |
| 21 | 2824 | 96.8 | 497 | 19 AAR69520 | rSTM protein SEQ I |
| 22 | 2820 | 96.7 | 497 | 17 AAR94607 | Human recombinant |
| 23 | 2803 | 96.1 | 494 | 21 AAY67401 | Novel sugar chain- |
| 24 | 2803 | 96.1 | 500 | 21 AAY69530 | Human thrombomodul |
| 25 | 2770 | 95.0 | 494 | 16 AAR78726 | Thrombomodulin wit |
| 26 | 2768 | 94.9 | 494 | 13 AAR24400 | Recombinant thromb |
| 27 | 2753 | 94.4 | 494 | 16 AAR78727 | Thrombomodulin wit |
| 28 | 2748 | 94.2 | 486 | 12 AAR13877 | Thrombin-binding s |
| 29 | 2720 | 93.3 | 480 | 13 AAR22013 | Truncated human th |
| 30 | 2690 | 92.2 | 475 | 13 AAR22032 | Truncated human ur |
| 31 | 2684 | 92.0 | 476 | 16 AAR78725 | Mature thrombomod |
| 32 | 2684 | 92.0 | 476 | 16 AAR86376 | Modified thrombomo |
| 33 | 2682 | 92.0 | 476 | 16 AAR86377 | Modified thrombomo |
| 34 | 2657 | 91.1 | 535 | 9 AAR80641 | Sequence of human |
| 35 | 2619 | 89.8 | 462 | 15 AAR45336 | Thrombomodulin ana |
| 36 | 2615 | 89.7 | 462 | 15 AAR45337 | Thrombomodulin ana |
| 37 | 2615 | 89.7 | 462 | 15 AAR45338 | Thrombomodulin ana |
| 38 | 2598 | 89.1 | 462 | 15 AAR45339 | Thrombomodulin ana |
| 39 | 2596 | 89.0 | 462 | 15 AAR45350 | Thrombomodulin ana |
| 40 | 2591 | 88.9 | 462 | 15 AAR45343 | Thrombomodulin ana |
| 41 | 2572 | 88.2 | 462 | 15 AAR45349 | Thrombomodulin ana |
| 42 | 2571 | 88.2 | 462 | 15 AAR45353 | Thrombomodulin ana |
| 43 | 2566 | 88.0 | 462 | 15 AAR45347 | Thrombomodulin ana |
| 44 | 2565.5 | 88.0 | 461 | 15 AAR45335 | Thrombomodulin ana |
| 45 | 2561 | 87.8 | 462 | 15 AAR45346 | Thrombomodulin ana |

ALIGNMENTS

RESULT 1
AAR22016 *
ID AAR22016 standard; Protein; 516 AA.
XX
AC AAR22016;
XX
DT 03-JUL-1992 (first entry)
XX
DE Truncated human thrombomodulin encoded by plasmid pSV2TMD1.
XX
KW Thrombin binding site; blood clotting; TMD1 deleter.
XX
OS Homo sapiens.
XX
PN EP474273-A.
XX
PD 11-MAR-1992.
XX
PF 05-AUG-1991; 91EP-0202009.
XX
PR 03-AUG-1990; 90JP-0204978.
XX
PA (ASAH) ASahi KASEI KOGYO.
XX
Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;
WPI; 1992-081820/11.
XX
PT New polypeptide inhibits blood coagulation and platelet
aggregation - promotes thrombin catalysed protein C activation
for treating myocardial infarction, thrombosis, embolism, etc.
XX
PS Example 1; Page 17; 112pp; English.
XX

```
CC Plasmid M13mp19TMJ3 (see AAR22013) was used as template for
CC site-directed mutagenesis. A 177bp fragment was deleted using the
CC "TMd1 deleter" oligonucleotide to give plasmid M13TMd1 which encodes
CC the first 516 N-terminal amino acids of human thrombomodulin. Plasmid
CC M13TMd1 was completely digested with HindIII and BamHI and a TMd1
CC fragment of ca. 1700bp was isolated. The fragment was ligated to
CC HindIII- and BglII-cut plasmid pSV2-dhfr to give the recombinant
CC plasmid pSV2TMd1. See AAR22014-R22022 and AAQ25072.
XX
SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 13; Length 516;
Best Local Similarity 100.0%; Pred. No. 2.2e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVLAGALAGLGFAPAPAPPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 mlgvlvlgallaglgfpapaeppggsqcvchdcfalyppgpatflnasqicdglrghlm 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGPRGFOWVTGDNNTSYS 120
DB 61 tvrssvaadvvisllngdggvrrrlwlgqlppgcgdpkrlgprlgfowvtgdnntsys 120
QY 121 RWARLDINGAPLCGPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEPHFATCRPLAV 180
DB 121 rwarldingaplcgplcvavsaeatvpsepiweeqcevkadgflcefhfatorplav 180
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGQLMCTAPPAGVQGHWAREAP 240
DB 181 epgaaaasvitygtptfaargadfgalpvgssaaavaplgqlmctappgavqghwareap 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcgcehacnaipgarcqcpagaaalqadgrsctasatqscndlcehfcvnp 300
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNVDLVG 360
DB 301 dqpgsyscmcetgyrllaadqhrcevdvdcilepspcqrcvntqggfechcypnydlvg 360
QY 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
DB 361 ecvepvpdpcfranceyqcqplnqtslylcvcagfapiphephrcomfcnotacpadcdpn 420
QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGVCVCHNLFTFECICGPDPSALVRHIGTDC 480
DB 421 tqascepegyllddgfictdidecengfcgvcvchnlftfecicgpdpsalvrhigtgc 480
QY 481 DSGKVDGDSGSGEPPTPGSTLTPPAPVGLVHSG 516
DB 481 dsgkvdgdsdsgsgppptpgstltppapvglvhs 516

RESULT 2
AA09347
ID AA09347 standard; Protein; 516 AA.
XX
XX AC AA09347;
XX
XX DT 08-JUL-1999 (first entry)
XX
XX DE Human thrombomodulin SEQ ID NO:1.
XX
XX KW Human; thrombomodulin; aqueous parenteral solution; storage;
XX KW distribution; acute coronary syndrome; thrombosis; embolism;
XX KW diabetes.
XX
XX OS Homo sapiens.
XX
XX PN W09918994-A1.
XX
XX PD 22-APR-1999.
XX
XX
```

```
PF 13-OCT-1998; 98WO-JP04609.
XX
PR 11-NOV-1997; 97JP-0308523.
PR 15-OCT-1997; 97JP-0281659.
XX
PA (ASAH ) ASahi KASEI KOGYO KK.
PI Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;
XX
XX WPI; 1999-277444/23.
DR N-PSDB; AAX55879.
XX
PT Stable aqueous parenteral thrombomodulin solution - comprising
PT buffer and surfactant, useful for treating acute coronary syndrome,
PT thrombosis, embolism, and diabetes
XX
PS Claim 6; Page 87-89; 97pp; Japanese.
XX
XX The present invention describes a method for maintaining the quality of
XX an aqueous parenteral solution of thrombomodulin comprising buffer and
XX surfactant aseptically filled in a case or syringe. Maintaining the
XX quality of an aqueous, parenteral thrombomodulin solution is
XX characterised in that the solution: (a) comprises soluble thrombomodulin,
XX a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically
XX filled into: (i) a case or (ii) a syringe without any empty space; and
XX (c) is kept in liquid form in storage and distribution and not frozen or
XX freeze dried. Thrombomodulin is used to treat acute coronary syndrome
XX (such as myocardial infarction, unstable angina and coronary artery
XX vessel thrombosis (e.g. cerebral, vascular and peripheral blood
XX vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.
XX Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic
XX lupus erythematosus or Barret's syndrome), multiple organ failure,
XX disseminated intravascular coagulation, transient ischaemia, diabetes,
XX liver veno-occlusive diseases and deep vein thrombosis. The composition
XX is stable for a long period of time and can be stored and distributed in
XX ready to use form avoiding the problems of dissolution and accuracy when
XX preparing on demand. The present sequence represents human
XX thrombomodulin.
XX
SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 20; Length 516;
Best Local Similarity 100.0%; Pred. No. 2.2e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVLAGALAGLGFAPAPAPPGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 mlgvlvlgallaglgfpapaeppggsqcvchdcfalyppgpatflnasqicdglrghlm 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPPGCGDPKRLGPRGFOWVTGDNNTSYS 120
DB 61 tvrssvaadvvisllngdggvrrrlwlgqlppgcgdpkrlgprlgfowvtgdnntsys 120
QY 121 RWARLDINGAPLCGPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEPHFATCRPLAV 180
DB 121 rwarldingaplcgplcvavsaeatvpsepiweeqcevkadgflcefhfatorplav 180
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGQLMCTAPPAGVQGHWAREAP 240
DB 181 epgaaaasvitygtptfaargadfgalpvgssaaavaplgqlmctappgavqghwareap 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcgcehacnaipgarcqcpagaaalqadgrsctasatqscndlcehfcvnp 300
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNVDLVG 360
DB 301 dqpgsyscmcetgyrllaadqhrcevdvdcilepspcqrcvntqggfechcypnydlvg 360
QY 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
DB 361 ecvepvpdpcfranceyqcqplnqtslylcvcagfapiphephrcomfcnotacpadcdpn 420
```

QY 421 TQASCEPEGYILLDDGFICTDIDECENGRCFSGVCHNLPTFTFCICGPDLSALVRHIGTDC 480
DB 421 tqascecegyillddgfictdidecengrcfsgvchnlptftfcicgpdlsalvrhigtcd 480
QY 481 DSGKVDGDSGSGEPPSPPTGTLTPPAVGLVHSG 516
DB 481 dschkvdgdsdgsgeppspptgsltppavglvhsq 516

RESULT 3
AAV83934
ID AAY83934 standard; Protein; 516 AA.
AC AAY83934;
XX 28-JUL-2000 (first entry)
XX Human thrombomodulin TMD protein.
DE Human; thrombomodulin; vasculitis; protein C; thrombin.
KW Homo sapiens.
OS JP2000053582-A.
PN 22-FEB-2000-
XX 06-AUG-1998; 98JP-0222688.
XX 06-AUG-1998; 98JP-0222688.
XX (ASAH) ASAH KASEI KOGYO KK.
XX WPI: 2000-353249/31.
XX N-PSDB; AAA10027.
XX Treating agent for vasculitis contains peptide which promotes
PT activation of protein C by thrombin -
XX Claim 4; Page 10-12; 18pp; Japanese.
XX This sequence represents a human thrombomodulin protein. The invention
CC relates to a treating agent for vasculitis containing a peptide which
CC promotes activation of protein C by thrombin.
XX Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 21; Length 516;
Best Local Similarity 100.0%; Pred. No. 2.2e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVIGALALAGLGFAPAEPQPGSGQVHDCFALYPGPATFLNASQICDGLRGLM 60
DB 1 mlgvlvigalalaglgfapaepqpgsgqvehdcfalyppatflnasqicdglrghlm 60
QY 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPPCGDPRKLRGFGQWVTGDNNTSYS 120
DB 61 tvrssvaadvvisllngdggvgrrrlwiglqppcgdprrklrgfgqwvtgdnntsys 120
QY 121 RWARLDLNGAPLCGPLCVAVSAAATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 180
DB 121 rwarldlngaplcgplcvavsaaatvpsepiweeqcevkadgflcefhfpatcrplav 180
QY 181 EPGAAAASVTYCTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
DB 181 epgaaaasvtyctpfaargadfoalpvgssaaavaplglqlmctappagvqghwareap 240
QY 241 GAWDCSVENGCGEACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcgacnaipgaprcqcpagaaqadgrsctasatqscndlcehfcvnp 300

QY 301 DQPGSYSCMETGYRLAADQHRCEVDYDCILERSPCQRCVNTQGGFEGCHCYPNDVLVDG 360
DB 301 dqpgsyscmetgyrllaadqhrcevdvdcilerspqrvcvntqggfegchcypnydlvdg 360
QY 361 ECVEPYDPCFRANCEYQCQPLNQTSLYLCVACGFAPIPHPHRCQFNCQTACPADCDPN 420
DB 361 ecvepydpcfranceyqcqplnqtsylcvacgfaiphephrcqfncqtacpadcdpn 420
QY 421 TQASCEPEGYILLDDGFICTDIDECENGRCFSGVCHNLPTFTFCICGPDLSALVRHIGTDC 480
DB 421 tqascecegyillddgfictdidecengrcfsgvchnlptftfcicgpdlsalvrhigtcd 480
QY 481 DSGKVDGDSGSGEPPSPPTGTLTPPAVGLVHSG 516
DB 481 dschkvdgdsdgsgeppspptgsltppavglvhsq 516

RESULT 4
AAV69529
ID AAY69529 standard; Protein; 516 AA.
AC AAY69529;
XX 10-APR-2000 (first entry)
XX Human thrombomodulin variant, SEQ ID NO:3.
DE Thrombomodulin; TM; soluble; affinity purification; cation exchange;
KW antioagulant; thrombolytic.
XX Homo sapiens.
XX JP11341990-A.
XX 14-DEC-1999.
XX 23-MAR-1999; 99JP-0077518.
XX 30-MAR-1998; 98JP-0084389.
XX (ASAH) ASAH KASEI KOGYO KK.
XX WPI: 2000-101696/09.
XX N-PSDB; AAZ55965.
XX Preparation of highly pure soluble thrombomodulin - used as an
PT antiblood coagulation agent and a thrombolytic agent
XX Claim 9; Page 30-32; 36pp; Japanese.
XX The invention relates to a novel method for the preparation of highly
CC pure soluble thrombomodulin (TM) containing substantially no serum-
CC derived or antibody-derived substance. The method comprises isolating
CC soluble TM from unpurified serum via affinity purification using an
CC anti-TM antibody. The soluble TM is then treated with a cation
CC exchanger at a specific conductivity of 25 to 34 ms/cm and a pH of 3 to
CC 4, and the fraction containing the soluble TM is isolated. The soluble
CC TM can be used as a blood anticoagulation agent and a thrombolytic agent.
CC This sequence represents a human thrombomodulin variant, designated
XX SEQ ID NO:3.
XX Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 21; Length 516;
Best Local Similarity 100.0%; Pred. No. 2.2e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVIGALALAGLGFAPAEPQPGSGQVHDCFALYPGPATFLNASQICDGLRGLM 60
DB 1 mlgvlvigalalaglgfapaepqpgsgqvehdcfalyppatflnasqicdglrghlm 60
QY 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPPCGDPRKLRGFGQWVTGDNNTSYS 120

Db 61 tvrsvaadvisllngdgvgrrrlwlgqlppgcgpkrlgplrgfqwvtgdnntsys 120
Qy 121 RWARLDNGALGCLPLCVAVSAATVPSEPIWEEQQCEVKADGFLCEHFPATCRPLAV 180
Db 121 rwarldngalgcplcvavsaatvpsepiweeqcevkdagdficefhfpatcrplav 180
Qy 181 EPGAAAAVSTYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 epgaavaavstygtpfaargadfqalpvgssaavaplgqlmctappgavqghwareap 240
Qy 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 gawdcsvengcehacnaipgprcqcpgaalqadgrsctasatqscndlcehfcvnp 300
Qy 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNYDLVDG 360
Db 301 dqpgyscmctgyrllaadqhrcevdvdcillepspcprcvntggfchecyptynydlvdg 360
Qy 361 ECVEPVPDFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ecvepvpdcfranceyqcqplnqtsylcvcaegfapiphephrcqmfcnqtacpadcdpn 420
Qy 421 TQASCECEGYILDDGFICTDIDECENGFGCGVCHNLPGTFECICGPDLSALVRHIGTDC 480
Db 421 tqascecegyilddgfictdidecengfgcgvchnlpgtfecicgpdlsalvrhigtcd 480
Qy 481 DSGKVDGDSGSGEPPPTGSLTTPPAVGLVHSG 516
Db 481 dschkvdgdsdsgseppptgstlttppavglvhsq 516

RESULT 5

AAR11534
ID AAR11534 standard; Protein; 575 AA.
XX AAR11534;
XX 17-JUN-1991 (first entry)
XX Human thrombomodulin type II polypeptide.
XX Thrombomodulin; chondroitin; protein C; thrombin; thrombolysis;
KW chondroitin sulphate; anticoagulant; myocardial infarction.
XX
FH Key Location/Qualifiers
FT - Peptide 1..18
FT - Peptide /label= signal peptide
FT Peptide 19..365
FT Peptide /label= opt. N-terminal sequence
FT Peptide 366..480
FT Peptide /label= human thrombomodulin active fragment
FT Peptide 481..516
FT Peptide /label= opt. C-terminal sequence
XX
XX WO9104276-A.
XX
XX
XX PD 04-APR-1991.
XX PF 25-SEP-1990; 90WO-JP01234.
XX PR 25-SEP-1989; 89JP-0246270.
XX (ASAH) ASAH KASEI KOGYO.
XX Yamamoto S, Gomi K, Ogawa K;
XX WPI; 1991-117478/16.
XX New human thrombomodulin polypeptide contg. chondroitin - has
PT anticoagulant platelet agglutination and thrombolytic activity
XX Disclosure; fig 1; 80pp; Japanese.
PS

XX This human thrombomodulin polypeptide comprises a 115 residue active
CC fragment with opt. flanking sequences (N-terminal= 347 amino acids;
CC C-terminal= 36 amino acids). Additionally it has an N-terminal signal
CC sequence and an attached sugar chain, contg. chondroitin or chondroitin
CC sulphate. It promotes the activation of protein C by thrombin and has
CC good anticoagulant, platelet aggregation inhibition and thrombolytic
CC activities. It can therefore be used for treating cardiovascular
CC diseases, eg arteriosclerosis or myocardial infarction.
XX Sequence 575 AA;

Query Match 100.0%; Score 2916; DB 12; Length 575;
Best Local Similarity 100.0%; Pred. No. 2.4e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLGVLVGALALAGLGFAPAEPPQPGSQCEVHDFCALYPGATFLNASQICDGLRHLM 60
Db 1 mlgvlvlgalalaglgfpapaeppqpgsqcevhdcfalyppatflnasqicdglrhlml 60
Qy 61 TVRSSVAADVISLLNGDGGVGRRLWLGQLPPCGDPPKRLGPLRGFQWVTGDNNTSYS 120
Db 61 tvrssvaadvlsllngdggvgrrrlwlgqlppcgdpkrlgplrgfqwvtgdnntsys 120
Qy 121 RWARLDNGALGCLPLCVAVSAATVPSEPIWEEQQCEVKADGFLCEHFPATCRPLAV 180
Db 121 rwarldngalgcplcvavsaatvpsepiweeqcevkdagdficefhfpatcrplav 180
Qy 181 EPGAAAAVSTYGTFFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 epgaavaavstygtpfaargadfqalpvgssaavaplgqlmctappgavqghwareap 240
Qy 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 gawdcsvengcehacnaipgprcqcpgaalqadgrsctasatqscndlcehfcvnp 300
Qy 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNYDLVDG 360
Db 301 dqpgyscmctgyrllaadqhrcevdvdcillepspcprcvntggfchecyptynydlvdg 360
Qy 361 ECVEPVPDFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
Db 361 ecvepvpdcfranceyqcqplnqtsylcvcaegfapiphephrcqmfcnqtacpadcdpn 420
Qy 421 TQASCECEGYILDDGFICTDIDECENGFGCGVCHNLPGTFECICGPDLSALVRHIGTDC 480
Db 421 tqascecegyilddgfictdidecengfgcgvchnlpgtfecicgpdlsalvrhigtcd 480
Qy 481 DSGKVDGDSGSGEPPPTGSLTTPPAVGLVHSG 516
Db 481 dschkvdgdsdsgseppptgstlttppavglvhsq 516

RESULT 6

AAR41806
ID AAR41806 standard; peptide; 575 AA.
XX AAR41806;
XX AC AAR41806;
XX 30-MAR-1994 (first entry)
XX Thrombomodulin.
XX Transformation; fungus; blood coagulation; prevention; platelet;
KW aggregation; thrombolytic activity; thrombolysis; human thrombomodulin;
KW site-directed mutagenesis; promotion; protein C; activation; thrombin.
XX Homo sapiens.
XX JP05213998-A.
XX 24-AUG-1993.
PD

```

XX 02-AUG-1991; 91JP-0282369.
XX
XX 03-AUG-1990; 90JP-0204978.
PR 30-JUL-1991; 91JP-0189984.
XX
XX (ASAH ) ASahi CHEM IND CO LTD.
XX
XX WPI; 1993-299652/38.
XX
XX Novel polypeptide obtd. by culturing transformed fungus - having
PT blood coagulation preventing, platelet aggregation preventing and
PT thrombolytic activities
XX
XX Disclosure; Fig 55; 65pp; Japanese.
XX
XX Novel polypeptides, obtd. by culturing transformed fungus, have
CC blood coagulation preventing, platelet aggregation preventing
CC and thrombolytic activities
CC
CC In an example, plasmid M13mpl19TMD3 (constructed from pSV2TMD2
CC (ATCC 67238) contg. a human thrombomodulin sequence) was subjected
CC to site directed mutagenesis to prepare pSV2TMD7. plasmid pSV2TMD7
CC was transfected to COS-1 cells. The activity of promoting protein C
CC activation by thrombin of the peptide produced by the transformed
CC COS-1 cell was measured. The amt. of the peptide was determined.
XX
XX Sequence 575 AA;
SQ
Query Match 100.0%; Score 2916; DB 14; Length 575;
Best Local Similarity 100.0%; Pred. No. 2.4e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLVVLGALALAGLGGFPAPAEPPQGGSCQVEHDCFALYPGPAPFLNASQICDGLRGHLM 60
DB 1 mlgvlvgalalaglgfpapaeppgsgqcvchdcfalyppgpatflnasqicdglrghlm 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWGLQLPPCGDPRKRLGPRGFWGTGDNNTSYS 120
DB 61 tvrssvaadvvisllngdggvrrrlwiglqlppcgdpkrllgprgfwgtgdnntsys 120
QY 121 RWARLDINGAPLGPCLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAV 180
DB 121 rwarldingaplgpclcvavsaeatvpsepieweeqcevkdgflcehfhpacrpplav 180
QY 181 EPGAASAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
DB 181 epgaasaaavstytgtptfaargadfoalpvgsaaavaplgqlmctappgavqghwareap 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQLQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcgcehacnaipgarcqcpagaaqlqadgrsctasatqscndlcehfcvnp 300
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYPNYDLVDG 360
DB 301 dqpgsyscmctgyrlaadqhrcevdvdcilepspqrcvntqggftechchypnydlvdg 360
QY 361 ECVEPVDPFRANCEYOCQPLNQTSYLCVCAEGFAPIPHEPHRCOMCNQACPADCDPN 420
DB 361 ecvepvdpfranceyocqplnqtsylcvcaegfapiphephrcomcnqacpadcdpn 420
QY 421 TOASCEPEGYILDGFGTCTDIDECENGFGSGVCHNLPGTFECICGPDLSALVRHIGTDC 480
DB 421 tqascepegyildgfgtctdidcengfgsgvchnlpgtfecicgpdlsalvrhigtcd 480
QY 481 DSGKVDGSDSGSGPPSPPTGSLTTPPAVGLVHSG 516
DB 481 dsgkvdgsgsgsgppspptgslttpavglvhs 516
RESULT 7
AAR43031
ID AAR43031 standard; protein; 575 AA.

```

```

XX AAR43031;
AC
XX
XX 16-MAY-1994 (first entry)
XX
XX Human thrombomodulin.
XX
XX Anticoagulant; platelet aggregation inhibitor;
KW protein C; activation; thrombin; thrombomodulin;
KW coagulation disorder; thrombosis; myocardial infarction;
KW embolism; telangiectasis; arteriosclerosis obliterans;
KW disseminated intravascular coagulation; DIC; angina pectoris;
KW gestosis; transient ischaemic attack.
XX
XX Homo sapiens.
XX
XX WO9322447-A.
XX
XX 11-NOV-1993.
XX
XX 30-APR-1993; 93WO-JP00578.
XX
XX 01-MAY-1992; 92JP-0112903.
XX
XX (ASAH ) ASahi CHEM IND CO LTD.
XX (ASAH ) ASahi KASEI KOGYO KK.
XX
XX Kondo S, Toma K, Zushi M;
XX
XX WPI; 1993-368806/46.
XX
XX Peptide with anticoagulant and platelet aggregation inhibitor
PT activity - which promotes protein C activation by thrombin and is
PT useful in treating coagulation disorders e.g. thrombosis
XX
XX Disclosure; Fig 1; 84pp; Japanese.
XX
XX New peptides (see AAR50069) are inhibitors of the blood coagulation
CC and platelet aggregation activities of thrombin and promote the
CC protein-C activation effect of thrombin. They can be produced
CC efficiently in pure form by culture of appropriate transformants,
CC and are useful in treatment of circulatory disorders such as
CC myocardial infarction, thrombosis, embolism, telangiectasis,
CC arteriosclerosis obliterans, disseminated intravascular
CC coagulation, angina pectoris, gestosis and transient ischaemic
CC attack.
XX
XX Sequence 575 AA;
SQ
Query Match 100.0%; Score 2916; DB 14; Length 575;
Best Local Similarity 100.0%; Pred. No. 2.4e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLVVLGALALAGLGGFPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 mlgvlvgalalaglgfpapaeppgsgqcvchdcfalyppgpatflnasqicdglrghlm 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWGLQLPPCGDPRKRLGPRGFWGTGDNNTSYS 120
DB 61 tvrssvaadvvisllngdggvrrrlwiglqlppcgdpkrllgprgfwgtgdnntsys 120
QY 121 RWARLDINGAPLGPCLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAV 180
DB 121 rwarldingaplgpclcvavsaeatvpsepieweeqcevkdgflcehfhpacrpplav 180
QY 181 EPGAASAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
DB 181 epgaasaaavstytgtptfaargadfoalpvgsaaavaplgqlmctappgavqghwareap 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQLQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcgcehacnaipgarcqcpagaaqlqadgrsctasatqscndlcehfcvnp 300

```

```

QY 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTGGFECHECHYPNYDLVDG 360
Db 301 dpgsyscmctgyrllaadqhrcevdvdcilep spqrcvntggf echechypnydlvdg 360
QY 361 ECVEPVPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQTFACADCDPN 420
Db 361 ecvepvpdpcfranceyqcqplnqtsylcvcaegfapiphephrcqmfcnqtacpadcdpn 420
QY 421 TQASCECEGYILDDGFICTDIDCECGGFCGVCCHNLPGTFECICGPDLSALVRHIGTDC 480
Db 421 tqascecegyilddgfictdidecengfcsgvchnlpgtfecicgpdlsalvrhigtcd 480
QY 481 DSGKVDGSGSGEPSPPTGSLTPPAVGLVHSG 516
Db 481 dsgkvdgsgsgseppspptgsltpppavglvhs g 516

RESULT 8
AAR22018
ID AAR22018 standard; Protein; 516 AA.
XX
AC AAR22018;
XX
DT 03-JUL-1992 (first entry)
XX
DE Human thrombomodulin (1-516) with Asp367 substituted by Glu.
XX
KW Mutant; thrombin binding site; blood clotting; TMm2 mutator.
XX
OS Homo sapiens.
XX
PN EP474273-A.
XX
PD 11-MAR-1992.
XX
PF 05-AUG-1991; 91EP-0202009.
XX
PR 03-AUG-1990; 90JP-0204978.
XX
PA (ASAH ) ASahi KASEI KOGYO.
XX
PI Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;
XX
DR WPI; 1992-081820/11.
XX
PT New polypeptide inhibits blood coagulation and platelet
PT aggregation - promotes thrombin catalysed protein C activation
PT for treating myocardial infarction, thrombosis, embolism, etc.
XX
PS Example 1; Page 18; 112pp; English.
XX
CC Plasmid M13TMD1 (see AAR22016) encoding the first 516 N-terminal
CC amino acids of human thrombomodulin was used as a template for
CC site-directed mutagenesis using the "TMm2 mutator" to produce
CC plasmid M13TMm2. In the mutant thrombomodulin encoded by the
CC plasmid, the wild-type Asp residue at position 367 is substituted
CC by an Glu residue. The activity of this mutant, truncated
CC thrombomodulin (i.e. D123Glu) in the activation of protein C was
CC compared to that of similarly truncated thrombomodulin with Asp
CC at position 367 (i.e. D123Asp). The activity of D123Glu was as high
CC as 1.2 times that of D123Asp. See AAR22013-R22022 and AAQ25074.
XX
SQ Sequence 516 AA;

```

```

Query Match 99.9%; Score 2012; DB 13; Length 516;
Best Local Similarity 99.8%; Pred. No. 3.7e-162;
Matches 515; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLCGLVLAGALAGLPAPAEPPGSGSCVCHNLPATFLNASQICDGLRGLM 60
Db 1 mlgvlgvalaglaglpapaeppgsgscvchnlpgatflnasqicdgllrghlm 60

```

```

QY 61 TVRSVAADVISLLNGDGVGRRRLWIGLQLPCCGDKRLGLRGFQWVTGDNNTSYS 120
Db 61 tvrsvaadvvisllngdgvgrrrllwiglqlppcgdkpklrglrgfqwtgdnntsys 120
QY 121 RWARDLNGAPLCGPLCVAVSAAEATVSEPIWEEQOCCEVKADGFLCFHFHPATCRPLAV 180
Db 121 rwardlngaplcgplcvavsaaeatvsepiweeqoccevkadgflcfhfhpattrplav 180
QY 181 EPGAAAAVSYTYGTPPAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVOGHWAREAP 240
Db 181 epgaaaaavsytygtppaargadfoalpvgssaaavaplglqlmctappagvoghwareap 240
QY 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFVCVNP 300
Db 241 gawdcsvenggehacnaipgprcqcpagaa lqadgrsctasatqscndlcehfcvnp 300
QY 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTGGFECHECHYPNYDLVDG 360
Db 301 dpgsyscmctgyrllaadqhrcevdvdcilep spqrcvntggf echechypnydlvdg 360
QY 361 ECVEPVPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQTFACADCDPN 420
Db 361 ecvepvpdpcfranceyqcqplnqtsylcvcaegfapiphephrcqmfcnqtacpadcdpn 420
QY 421 TQASCECEGYILDDGFICTDIDCECGGFCGVCCHNLPGTFECICGPDLSALVRHIGTDC 480
Db 421 tqascecegyilddgfictdidecengfcsgvchnlpgtfecicgpdlsalvrhigtcd 480
QY 481 DSGKVDGSGSGEPSPPTGSLTPPAVGLVHSG 516
Db 481 dsgkvdgsgsgseppspptgsltpppavglvhs g 516

```

RESULT 9

```

AAY09348
ID AAY09348 standard; Protein; 516 AA.
XX
AC AAY09348;
XX
DT 08-JUL-1999 (first entry)
XX
DE Human thrombomodulin SEQ ID NO:2.
XX
KW Human; thrombomodulin; aqueous parenteral solution; storage;
KW distribution; acute coronary syndrome; thrombosis; embolism;
XX diabetes.
XX
OS Homo sapiens.
XX
PN WO9918994-A1.
XX
PD 22-APR-1999.
XX
PF 13-OCT-1998; 98WO-JP04609.
XX
PR 11-NOV-1997; 97JP-0308523.
PR 15-OCT-1997; 97JP-0281659.
XX
PA (ASAH ) ASahi KASEI KOGYO KK.
XX
PI Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;
XX
DR WPI; 1999-277444/23.
DR N-PSDB; AAX55880.
XX

```

```

PT Stable aqueous parenteral thrombomodulin solution - comprising
PT buffer and surfactant, useful for treating acute coronary syndrome,
PT thrombosis, embolism, and diabetes
XX
PS Claim 7; Page 90-92; 97pp; Japanese.
XX

```

CC The present invention describes a method for maintaining the quality of

an aqueous parenteral solution of thrombomodulin comprising buffer and surfactant aseptically filled in a case or syringe. Maintaining the quality of an aqueous, parenteral thrombomodulin solution is characterised in that the solution: (a) comprises soluble thrombomodulin, a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically filled into: (i) a case or (ii) a syringe without any empty space; and (c) is kept in liquid form in storage and distribution and not frozen or freeze dried. Thrombomodulin is used to treat acute coronary syndrome (such as myocardial infarction, unstable angina and coronary artery blockage), thrombosis (e.g. cerebral, vascular and peripheral blood vessel thrombosis), embolism, peripheral blood vessel disorders (e.g. Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic lupus erythematosus or Barrett's syndrome), multiple organ failure, disseminated intravascular coagulation, transient ischaemia, diabetes, liver veno-occlusive diseases and deep vein thrombosis. The composition is stable for a long period of time and can be stored and distributed in ready to use form avoiding the problems of dissolution and accuracy when preparing on demand. The present sequence represents human thrombomodulin.

XX Sequence 516 AA;

Query Match 99.9%; Score 2912; DB 20; Length 516;
Best Local Similarity 99.8%; Pred. No. 3.7e-162;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLVLGALAGLGFPAAPPGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60
DB 1 mlgvlgvlgalaaglgfpapepgpgsgqvehdcfalyppgpatflnasqicdglrghlm 60
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWGLQLPGCGDPKRLGPRGQVWTDGNTSYS 120
DB 61 tvrssvaadvlsllngdggvrrrlwglqlpgcgdpkrlgprgqvwtdgntsys 120
QY 121 RWARLDNGAPLCGPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 180
DB 121 rwarldngaplcgplcvavsaeeatvpsepiweeqcevkadgflcfefhpatcrplav 180
QY 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLOLMCTAPPAGVQGHWAREAP 240
DB 181 epgaaaaavstygtpfaargadfqalpvgssaaavaplglolmctappagvqghwareap 240
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcgchacnaipgaprcqcpagaaalqadgrsctasatqscndlcehfcvnp 300
QY 301 DPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYPNYDLVDG 360
DB 301 dpgsyscmctgyrlaadqhrcevdvdcilepspcrcvntqggfchecychnydlvdg 360
QY 361 ECVEPVPDFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
DB 361 ecvepvpdfranceyqcqplnqtsylcvcaegfapiphephrcmfncqtacpadcdpn 420
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPFTGTCIGCPDPSALVRHIGTDC 480
DB 421 tqascecegyilddgfictdidecengfcgvcvchnlpftgfcicgpdpsalarhigtcd 480
QY 481 DSGKVDGSGSGGEPPTGCTLTTPPAVGLVHSG 516
DB 481 dsgkvdgsgsggeppptgstltppavglvhs 516

RESULT 10

AA83935

ID AA83935 standard; Protein; 516 AA.

XX

AC AA83935;

XX

DT 28-JUL-2000 (first entry)

XX

DE Human thrombomodulin TMD protein #2.

XX Human; thrombomodulin; vasculitis; protein C; thrombin.
XX Homo sapiens.
XX JP2000053582-A.
XX 22-FEB-2000.
XX 06-AUG-1998; 98JP-0222688.
XX 06-AUG-1998; 98JP-0222688.
XX (ASAH) ASahi KASEI KOGYO KK.
XX WPI; 2000-353249/31.
XX N-PSDB; AAA10028.
XX Treating agent for vasculitis contains peptide which promotes activation of protein C by thrombin -
XX Claim 4; Page 13-14; 18pp; Japanese.
XX This sequence represents a human thrombomodulin protein. The invention relates to a treating agent for vasculitis containing a peptide which promotes activation of protein C by thrombin.
XX Sequence 516 AA;

Query Match 99.9%; Score 2912; DB 21; Length 516;
Best Local Similarity 99.8%; Pred. No. 3.7e-162;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVLVLGALAGLGFPAAPPGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60
DB 1 mlgvlgvlgalaaglgfpapepgpgsgqvehdcfalyppgpatflnasqicdglrghlm 60
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWGLQLPGCGDPKRLGPRGQVWTDGNTSYS 120
DB 61 tvrssvaadvlsllngdggvrrrlwglqlpgcgdpkrlgprgqvwtdgntsys 120
QY 121 RWARLDNGAPLCGPLCVAVSAEAATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 180
DB 121 rwarldngaplcgplcvavsaeeatvpsepiweeqcevkadgflcfefhpatcrplav 180
QY 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLOLMCTAPPAGVQGHWAREAP 240
DB 181 epgaaaaavstygtpfaargadfqalpvgssaaavaplglolmctappagvqghwareap 240
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 gawdcsvengcgchacnaipgaprcqcpagaaalqadgrsctasatqscndlcehfcvnp 300
QY 301 DPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYPNYDLVDG 360
DB 301 dpgsyscmctgyrlaadqhrcevdvdcilepspcrcvntqggfchecychnydlvdg 360
QY 361 ECVEPVPDFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
DB 361 ecvepvpdfranceyqcqplnqtsylcvcaegfapiphephrcmfncqtacpadcdpn 420
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPFTGTCIGCPDPSALVRHIGTDC 480
DB 421 tqascecegyilddgfictdidecengfcgvcvchnlpftgfcicgpdpsalarhigtcd 480
QY 481 DSGKVDGSGSGGEPPTGCTLTTPPAVGLVHSG 516
DB 481 dsgkvdgsgsggeppptgstltppavglvhs 516

RESULT 11

AA822017


```
FT Peptide 1..20
FT Domain /note= "signal peptide"
FT 245..280
FT /note= "EGF-1"
FT 288..323
FT /note= "EGF-2"
FT 329..361
FT /note= "EGF-3"
FT 368..404
FT /note= "EGF-4"
FT 408..439
FT /note= "EGF-5"
FT 445..480
FT /note= "EGF-6"
FT 481..515
FT /note= "O-linked glycosylation domain"
FT 516..538
FT Region /note= "stop transfer sequence"
FT 539..575
FT Domain /note= "cytoplasmic domain"
FT
XX WO9301282-A.
XX
XX 21-JAN-1993.
XX
XX 01-JUL-1992; 92WO-US05573.
XX
XX 01-JUL-1991; 91US-0724237.
XX
XX (BERL-) BERLEX LAB INC.
XX
XX Andrews WH, Morser MJ, Vilander LR;
XX
XX WPI; 1993-045488/05.
XX
XX Site-specific mutagenesis of nucleic acid constructs - using an
XX oligonucleotide which changes a nucleotide and introduces or
XX removes a restriction site
XX
XX Example; Fig 3; 87pp; English.
XX
XX The sequence is that of the native human thrombomodulin showing the
XX six EGF-like domains.
XX
XX Sequence 575 AA;
XX
XX
XX Query Match 99.6%; Score 2904; DB 14; Length 575;
XX Best Local Similarity 99.6%; Pred. No. 1.2e-161;
XX Matches 514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
Qy 1 MLGVLVLGALAGLGFAPAPGPGSQCVHEHDFCFALYGPATFLNASQICDGLRGHLM 60
Db 1 mlgvlvlgalalaglgfpapapgpqsgcvhehdcscalyppatflnasqicdglrghlm 60
Qy 61 TVRSVAADVLSLLNGDGGVRRRLWIGLQLPDGGDKRLGLRFGVQWVTGDNNTSYS 120
Db 61 tvrsvaadvlsllngdggvrrrlwlgldppgdcgdkrlglrgfvqwtgdnntsys 120
Qy 121 RWARLDLNGPLCGPLCVANSAEATVPSEPIWEEOQCEVKADGFLCEFFHPATCRPLAV 180
Db 121 rwarldngplcgpplcvansaeadvpsepiweeqqcevkadgflceffhpatcrplav 180
Qy 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 epgaaaaavstytptfaargadfoalpvgssaaavaplgqlmctappagvqghwareap 240
Qy 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 gawdcsvengcehacnaipgaprcqcpagaaalqadgrsctasatqscndlcehfcvnp 300
Qy 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTGGFECHCYPNYDLVDG 360
Db 301 dqpgsyscmctgyrllaadqhrcevdvdcilepspqrcvntggfchcypnydlvdg 360
```

```
Db 301 dqpgsyscmctgyrllaadqhrcevdvdcilepspqrcvntggfchcypnydlvdg 360
Qy 361 ECVEPVPDFCFRANCEYQCQLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420
Db 361 ecvepvpdfcfranceyqcqlnqtsylcvcaegfapiphephrcmqfncqtacpadcdpn 420
Qy 421 TOASCECEGYILDDGFICTDIDECENGCGFCGVCHNLPGTFECICGPDLSALVRHIGTDC 480
Db 421 tqascecegyilddgfictdidecengcgfcgvchnlpgtfecicgpdlsalarhigtcd 480
Qy 481 DSGKVDGDSGSGEPSPPTFGSLTPPAVGLVHSG 516
Db 481 dschkvdgdsdsgseppspptfgsltpavglvhsq 516
XX
XX RESULT 15
XX AAR20639
XX ID AAR20639 standard; Protein; 575 AA.
XX AC AAR20639;
XX XX
XX DT 01-MAY-1992 (first entry)
XX DE Human urinary thrombomodulin.
XX KW Blood clotting; intravascular coagulation.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 1..18
XX FT Protein /label= signal
XX FT /label= thrombomodulin
XX
XX WO9200325-A.
XX 09-JAN-1992.
XX
XX 27-JUN-1991; 91WO-JP00873.
XX
XX 27-JUN-1990; 90JP-0168766.
XX (MOCH ) MOCHIDA PHARM KK.
XX
XX Nii A, Morishita H, Uemura A, Mochida E;
XX
XX WPI; 1992-041517/05.
XX N-PSDB; AAQ20810.
XX
XX New modified recombinant human urinary thrombomodulin - has
XX thrombin binding, anticoagulant and thrombolytic activity and is
XX for treating blood coagulation disorders
XX
XX Claim 1; Fig 3; 107pp; Japanese.
XX
XX A 2.5kb cDNA fragment was isolated by screening a human placental
XX cell library with a probe coding for the N-terminal of human
XX thrombomodulin. This full-length sequence was modified to produce
XX two alternative, truncated forms of the coding sequence encoding
XX only up to amino acid 456 of the mature polypeptide; the Ala
XX residue at position 455 is substituted by Val in one of the two
XX truncated forms.
XX
XX Sequence 575 AA;
XX
XX
XX Query Match 99.4%; Score 2899; DB 13; Length 575;
XX Best Local Similarity 99.4%; Pred. No. 2.3e-161;
XX Matches 513; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
Qy 1 MLGVLVLGALAGLGFAPAPGPGSQCVHEHDFCFALYGPATFLNASQICDGLRGHLM 60
Db 1 mlgvlvlgalalaglgfpapapgpqsgcvhehdcscalyppatflnasqicdglrghlm 60
```

| | | | |
|----|-----|--|-----|
| Db | 1 | m1gvlvlgalalaglfpapaeppqpgsgqvehdcfalypgpatfinaaqicdglrgqlm | 60 |
| Qy | 61 | TVRSSVAADVISLLLLNGDGVGRRRLWIGLQLPPGCGDKRKLGLRGFOWVTDGNNNTSYS | 120 |
| Db | 61 | TVRSSVADVISLLLLNGDGVGRRRLWIGLQLPPGCGDKRKLGLRGFOWVTDGNNNTSYS | 120 |
| Qy | 121 | RWARLDLNGAPLCGPLICVAVSAEATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAY | 180 |
| Db | 121 | RWARLDLNGAPLCGPLICVAVSAEATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAY | 180 |
| Qy | 181 | EPGAAAAYSIYGTFFARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP | 240 |
| Db | 181 | EPGAAAAYSIYGTFFARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP | 240 |
| Qy | 241 | GAWDCSVENGGEHACNAIPGAPRCOCPAGALQADGRSCTASATOSCNDLCEHFVCPNP | 300 |
| Db | 241 | GAWDCSVENGGEHACNAIPGAPRCOCPAGALQADGRSCTASATOSCNDLCEHFVCPNP | 300 |
| Qy | 301 | DQPGSYSCMCETGYRLAABQHRCEVDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG | 360 |
| Db | 301 | DQPGSYSCMCETGYRLAABQHRCEVDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG | 360 |
| Qy | 361 | ECVEPVDPCFRANCEYQCOPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN | 420 |
| Db | 361 | ECVEPVDPCFRANCEYQCOPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN | 420 |
| Qy | 421 | TQASCECPGYILDDGFICTDDECENGGFCSCGVCHNLPGTFECICGPDSSALVRHIGTDC | 480 |
| Db | 421 | TQASCECPGYILDDGFICTDDECENGGFCSCGVCHNLPGTFECICGPDSSALVRHIGTDC | 480 |
| Qy | 481 | DSGKVDGDSGSGEPPPTPGSTLTPPAVGLVHSG | 516 |
| Db | 481 | DSGKVDGDSGSGEPPPTPGSTLTPPAVGLVHSG | 516 |

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:45:14 ; Search time 17.04 Seconds
(without alignments)
1110.273 Million cell updates/sec

Title: US-09-509-994-2
Perfect score: 2916
Sequence: 1 MLGVLVLCALALAGLGPAP.....PSPTPGSTLTTPPAVLVHSG 516

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query % | Length | DB | ID | Description |
|------------|--------|---------|--------|----|-----------------|---------------------|
| 1 | 2916 | 100.0 | 575 | 1 | TREML_HUMAN | P07204 homo sapien |
| 2 | 1903 | 65.3 | 577 | 1 | TREML_MOUSE | P15306 mus musculus |
| 3 | 1095.5 | 37.6 | 356 | 1 | TREML_BOVIN | P06579 bos taurus |
| 4 | 374 | 12.8 | 2907 | 1 | FBN2_MOUSE | Q61555 mus musculus |
| 5 | 368 | 12.6 | 2871 | 1 | FBN1_BOVIN | P98133 bos taurus |
| 6 | 367.5 | 12.6 | 1184 | 1 | FBN2_HUMAN | P98095 homo sapien |
| 7 | 366 | 12.6 | 2871 | 1 | FBN1_PIG | Q9TV36 sus scrofa |
| 8 | 366 | 12.6 | 2911 | 1 | FBN2_HUMAN | P35556 homo sapien |
| 9 | 364 | 12.5 | 2871 | 1 | FBN1_HUMAN | P35555 homo sapien |
| 10 | 362 | 12.4 | 2871 | 1 | FBN1_MOUSE | Q61554 mus musculus |
| 11 | 357 | 12.2 | 1221 | 1 | FBN2_MOUSE | P37889 mus musculus |
| 12 | 356 | 12.2 | 956 | 1 | MTN2_HUMAN | O00339 homo sapien |
| 13 | 332.5 | 11.4 | 1964 | 1 | NTC4_MOUSE | P31695 mus musculus |
| 14 | 330.5 | 11.3 | 1712 | 1 | TGFB_RAT | Q00918 rattus norv |
| 15 | 326.5 | 11.2 | 956 | 1 | MTN2_MOUSE | O08746 mus musculus |
| 16 | 324.5 | 11.1 | 712 | 1 | FBLL_CAEEL | O77469 caenorhabdi |
| 17 | 322.5 | 11.1 | 1394 | 1 | TGFB_HUMAN | P22064 homo sapien |
| 18 | 315.5 | 10.8 | 443 | 1 | FBLL_HUMAN | O95967 homo sapien |
| 19 | 311 | 10.7 | 703 | 1 | FBLL_MOUSE | P23142 homo sapien |
| 20 | 308.5 | 10.6 | 443 | 1 | FBLL_CRIGR | O55058 cricetus |
| 21 | 308 | 10.6 | 448 | 1 | FBLL_HUMAN | Q9UBX5 homo sapien |
| 22 | 308 | 10.6 | 705 | 1 | FBLL_MOUSE | Q08879 mus musculus |
| 23 | 307 | 10.5 | 684 | 1 | FBLL_CHICK | O73775 gallus gall |
| 24 | 306.5 | 10.5 | 443 | 1 | FBLL_MOUSE | Q9WVJ9 mus musculus |
| 25 | 302 | 10.4 | 448 | 1 | FBLL_MOUSE | Q9WVH9 mus musculus |
| 26 | 297 | 10.2 | 2531 | 1 | NTC1_MOUSE | Q01705 mus musculus |
| 27 | 296 | 10.2 | 448 | 1 | FBLL_RAT | Q9WVH8 rattus norv |
| 28 | 289.5 | 9.9 | 493 | 1 | FBLL_HUMAN | Q12805 homo sapien |
| 29 | 284.5 | 9.8 | 2437 | 1 | NOTC_BRARE | P46530 brachydanio |
| 30 | 282.5 | 9.7 | 2703 | 1 | NOTC_DROSOPHILA | P07207 drosophila |
| 31 | 280 | 9.6 | 2318 | 1 | NTC3_MOUSE | Q61982 mus musculus |
| 32 | 278 | 9.5 | 2531 | 1 | NTC1_RAT | Q07008 rattus norv |
| 33 | 276.5 | 9.5 | 493 | 1 | FBLL3_RAT | O35568 rattus norv |

| | | | | | | |
|----|-------|-----|------|---|------------|---------------------|
| 34 | 272 | 9.3 | 2444 | 1 | NTC1_HUMAN | P46531 homo sapien |
| 35 | 269.5 | 9.2 | 1217 | 1 | EGF_MOUSE | P01132 mus musculus |
| 36 | 263.5 | 9.0 | 619 | 1 | MTN4_HUMAN | O95460 homo sapien |
| 37 | 257.5 | 8.8 | 2524 | 1 | NOTC_XENLA | P21783 xenopus lae |
| 38 | 252.5 | 8.7 | 816 | 1 | NEL2_RAT | Q62918 rattus norv |
| 39 | 252.5 | 8.7 | 835 | 1 | NEL2_HUMAN | P48960 homo sapien |
| 40 | 251.5 | 8.6 | 816 | 1 | CD97_HUMAN | O99435 homo sapien |
| 41 | 248 | 8.5 | 624 | 1 | MTN4_MOUSE | O89029 mus musculus |
| 42 | 247.5 | 8.5 | 1133 | 1 | EGF_RAT | P07522 rattus norv |
| 43 | 247.5 | 8.5 | 3051 | 1 | YXN3_CAEEL | P34576 caenorhabdi |
| 44 | 247 | 8.5 | 816 | 1 | NEL2_MOUSE | Q61220 mus musculus |
| 45 | 247 | 8.5 | 816 | 1 | NEL_CHICK | Q90827 gallus gall |

ALIGNMENTS

| | |
|------------|--|
| RESULT 1 | |
| TRBM_HUMAN | |
| ID | TRBM_HUMAN |
| AC | P07204; |
| DT | 01-APR-1988 (Rel. 07, Created) |
| DT | 01-FEB-1991 (Rel. 17, Last sequence update) |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) |
| DE | THROMBOMODULIN PRECURSOR (FETOMODULIN) (TM) (CD141 ANTIGEN). |
| GN | THBD OR TRBM. |
| OS | Homo sapiens (Human). |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| OX | NCBI_TaxID=9606; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE=88004395; PubMed=2820710; |
| RA | Suzuki K., Kusumoto H., Deyashiki Y., Nishioka J., Maruyama I., |
| RA | Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.; |
| RT | "Structure and expression of human thrombomodulin, a thrombin |
| RT | receptor on endothelium acting as a cofactor for protein C |
| RT | activation."; |
| RL | EMBO J. 6:1891-1897(1987). |
| RN | [2] |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE=88024950; PubMed=2822087; |
| RA | Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E.; |
| RT | "Human thrombomodulin: complete cDNA sequence and chromosome |
| RT | localization of the gene."; |
| RL | Biochemistry 26:4350-4357(1987). |
| RN | [3] |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE=87317665; PubMed=2819876; |
| RA | Jackson R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D.; |
| RT | "Human thrombomodulin gene is intron depleted: nucleic acid sequences |
| RT | of the cDNA and gene predict protein structure and suggest sites of |
| RT | regulatory control."; |
| RL | Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987). |
| RN | [4] |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE=88227901; PubMed=2836377; |
| RA | Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H., |
| RA | Deyashiki Y., Maruyama I., Suzuki K.; |
| RT | "Gene structure of human thrombomodulin, a cofactor for thrombin- |
| RT | catalyzed activation of protein C."; |
| RL | J. Biochem. 103:281-285(1988). |
| RN | [5] |
| RP | SEQUENCE FROM N.A. |
| RA | Matthews L.; |
| RT | Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases. |
| RN | [6] |
| RP | CARBOHYDRATE-LINKAGE SITE SER-492, AND MUTAGENESIS. |
| RA | MEDLINE=94029900; PubMed=8216207; |
| RA | Gerleitz B., Hassell T., Vlahos C.J., Parkinson J.F., Bang N.U., |
| RA | Grinnell B.W.; |
| RT | "Identification of the predominant glycosaminoglycan-attachment site |
| RT | in soluble recombinant human thrombomodulin: potential regulation of |

RT functionality by glycosyltransferase competition for serine474.";
RL Biochem. J. 295:131-140(1993).
RN [7]
RN STRUCTURE BY NMR OF 389-407.
RX MEDLINE-96007474; PubMed-7559494;
RA Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.;
RT "The structure of a 19-residue fragment from the C-loop of the fourth
RT epidermal growth factor-like domain of thrombomodulin.";
RL J. Biol. Chem. 270:23366-23372(1995).
RN [9]
RN STRUCTURE BY NMR OF 364-407.
RX MEDLINE-96100636; PubMed-8528067;
RA Meininger D.P., Hunter M.J., Komives E.A.;
RT "Synthesis, activity, and preliminary structure of the fourth
RT EGF-like domain of thrombomodulin.";
RL Protein Sci. 4:1683-1695(1995).
RN [9]
RN STRUCTURE BY NMR OF 427-444.
RX MEDLINE-95034791; PubMed-7947766;
RA Srinivasan J., Hu S., Hrabal R., Zhu Y., Komives E.A., Ni F.;
RT "Thrombin-bound structure of an EGF subdomain from human
RT thrombomodulin determined by transferred nuclear Overhauser
RT effects.";
RL Biochemistry 33:13553-13560(1994).
RN [10]
RN STRUCTURE BY NMR OF 427-444.
RX MEDLINE-96276211; PubMed-8745396;
RA Hrabal R., Komives E.A., Ni F.;
RT "Structural resiliency of an EGF-like subdomain bound to its target
RT protein, thrombin.";
RL Protein Sci. 5:195-203(1996).
RN [11]
RN STRUCTURE BY NMR OF 405-444.
RX MEDLINE-98035729; PubMed-9367781;
RA Sampoli Benitez B.A., Hunter M.J., Meininger D.P., Komives E.A.;
RT "Structure of the fifth EGF-like domain of thrombomodulin: an
RT EGF-like domain with a novel disulfide-bonding pattern.";
RL J. Mol. Biol. 273:913-926(1997).
RN [12]
RN VARIANT TED TYR-486.
RX PubMed-7811989;
RA Oehlin A.-K., Marlar R.A.;
RT "The first mutation identified in the thrombomodulin gene in a
RT 45-year-old man presenting with thromboembolic disease.";
RL Blood 85:330-336(1995).
RN [13]
RN VARIANT TED Y-486, AND VARIANTS T-43; A-79; S-495 AND L-501.
RX PubMed-9198186;
RA Oehlin A.-K., Norlund L., Marlar R.A.;
RT "Thrombomodulin gene variations and thromboembolic disease.";
RL Thromb. Haemost. 78:396-400(1997).
RN [14]
RN VARIANT VAL-473.
RX PubMed-9157575;
RA Norlund L., Holm J., Zoller B., Oehlin A.-K.;
RT "A common thrombomodulin amino acid dimorphism is associated with
RT myocardial infarction.";
RL Thromb. Haemost. 77:248-251(1997).
RN [15]
RN VARIANT THR-43.
RX PubMed-9843165;
RA Doggen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,
RA Stubbs P.J., Manger Cats V., Ireland H.;
RT "A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr,
RT and the risk of myocardial infarction in men.";
RL Thromb. Haemost. 80:743-748(1998).
RN [16]
RN VARIANT VAL-473.
RX PubMed-11245641;
RA Wu K.K., Aleksic N., Ahn C., Boerwinkle E., Folsom A.R.,
RA Juneja H.;
RT "Thrombomodulin Ala455Val polymorphism and risk of coronary heart
RT disease.";

Circulation 103:1386-1389(2001).
RL
CC -1- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR
CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS
CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE
CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA
CC SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,
CC FACTOR VA AND FACTOR VIIIa, AND THEREBY REDUCES THE AMOUNT OF
CC THROMBIN GENERATED.
CC
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC
CC -1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING
CC THROMBOMODULIN.
CC
CC -1- POLYMORPHISM: VARIATIONS IN THBD ARE ASSOCIATED WITH AN INCREASED
CC RISK OF DEVELOPING THROMBOEMBOLIC DISEASES (TED).
CC
CC -1- DISEASE: DEFECTS IN THBD COULD BE THE CAUSE OF INHERITED TED,
CC ALSO CALLED INHERITED THROMBOPHILIA. PATIENTS WITH TED HAVE
CC DEFECTS OF THE HAEMOPOIETIC SYSTEM WHICH CREATES A TENDENCY TO THE
CC OCCURRENCE OF THROMBOSIS. TED PLAY AN IMPORTANT ROLE IN THE
CC PATHOGENESIS OF VARIOUS CARDIOVASCULAR DISORDERS.
CC
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC
CC -1- DATABASE: NAME-PRO; NOTE-CD guide CD141 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd141.htm".
CC
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation--
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X05495; CAA29045.1; -
CC EMBL; M16552; AAB59508.1; -
CC EMBL; J02973; AAA61175.1; -
CC EMBL; D00210; BAA00149.1; -
CC EMBL; AL049651; CAB51954.1; -
CC PIR; A27073; A27073.
CC PIR; A28307; A28307.
CC PIR; A29680; A29680.
CC PDB; 1EGT; 15-NOV-95.
CC PDB; 1FGD; 20-JUN-96.
CC PDB; 1FGE; 20-JUN-96.
CC PDB; 1TMR; 08-JUN-95.
CC PDB; 1ZAO; 29-JAN-96.
CC PDB; 1ADX; 24-DEC-97.
CC PDB; 2ADX; 24-DEC-97.
CC GlycoSuiteDB; P07204; -
CC MIN; 188040; -
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR001491; Thrombomoduln.
CC InterPro; IPR001304; lectin_c.
CC Pfam; PF00008; EGF; 5.
CC Pfam; PF00059; lectin_c; 1.
CC PRINTS; PR00907; THROMBOMODULN.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00001; EGF_like; 5.
CC PROSITE; PS00010; ASX_HYDROXYL; 2.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 2.
CC PROSITE; PS01187; EGF_CA; 2.
CC PROSITE; PS50041; C-TYPE LECTIN 2; 1.
CC Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;
KW Glycoprotein; Signal; EGF-like domain; Disease mutation; Polymorphism;
KW 3D-structure.
FT SIGNAL 1 21
FT CHAIN 22 575
FT DOMAIN 22 515
FT TRANSMEM 516 539
FT DOMAIN 540 575
FT DOMAIN 241 281
FT DOMAIN 284 324
FT POTENTIAL.
FT THROMBOMODULIN.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT EGF-LIKE 1.
FT EGF-LIKE 2.


```
FT CARBOHYD 494 494 O-LINKED (GLYCOSAMINOGLYCAN) (BY
FT SIMILARITY)
SQ SEQUENCE 577 AA; 61867 MW; B20E50B0FE745014 CRC64;

Query Match 65.3%; Score 1903; DB 1; Length 577;
Best Local Similarity 65.6%; Pred. No. 2.1e-118;
Matches 342; Conservative 46; Mismatches 125; Indels 8; Gaps 4;

QY 1 MLGVLVLGALALAGLFPAPAPQPGSQCVHEHDFCFALYPGPAPFLNASQICDGLRGHLM 60
DB 1 MLGIFLVLGAPASLGLSALAKLQPTGSCQVHEHDFCFALYPGPAPFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDG-GVGRRRRLWIGLQPPGCGDKPKRLGRLGFGQWVTGDNNTSY 119
DB 61 TVRSSVAADVISLLNGDGLG- --WIGLQPPGCGDKPKRLGRLGFGQWVTGDNNTSY 117
QY 120 SRWRLDNLGAPLCGLVAVSAEAATVPSEPTWEEQQCEVKADGFLCEPHFPATCRPLA 179
DB 120 SRWRLDNLGAPLCGLVAVSAEAATVPSEPTWEEQQCEVKADGFLCEPHFPATCRPLA 179
QY 118 SRWARPNDQAPLGGPLCVTVSTATAPGAEPAAWEEKPCETETQGFLCEFYFTASCRPLT 177
DB 118 SRWARPNDQAPLGGPLCVTVSTATAPGAEPAAWEEKPCETETQGFLCEFYFTASCRPLT 177
QY 180 VEP-GAAAAAVSYTYGPFPAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWARE 238
DB 180 VEP-GAAAAAVSYTYGPFPAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWARE 238
QY 178 VNTDPPEAAHISSTYNTYFPGVSGADFOTLPVGSAAVEPLGLELVCRAPPGTSEGHWARE 237
DB 178 VNTDPPEAAHISSTYNTYFPGVSGADFOTLPVGSAAVEPLGLELVCRAPPGTSEGHWARE 237
QY 239 APGAWDCSVENGCGEHAACNAIPGAPRCQCPAGAAALQADGRSCATASATQSCNDLCEHFVCP 298
DB 239 APGAWDCSVENGCGEHAACNAIPGAPRCQCPAGAAALQADGRSCATASATQSCNDLCEHFVCP 298
QY 238 ATGAWNCVSVENGCGEYLCNRSNEPRCLCPDMDLQADGRSCARPPVVSQCNELCEHFVCS 297
DB 238 ATGAWNCVSVENGCGEYLCNRSNEPRCLCPDMDLQADGRSCARPPVVSQCNELCEHFVCS 297
QY 299 NPQPGSYSCMCTGTGYRLADQHRCEVDVDCILEPSPCQPCQRCVNTQGGFCHCYPNVDLV 358
DB 299 NPQPGSYSCMCTGTGYRLADQHRCEVDVDCILEPSPCQPCQRCVNTQGGFCHCYPNVDLV 358
QY 298 NAEVPGSYSCMCTGTGYRLADQHRCEVDVDCILEPSPCQPCQRCVNTQGGFCHCYPNVDLV 357
DB 298 NAEVPGSYSCMCTGTGYRLADQHRCEVDVDCILEPSPCQPCQRCVNTQGGFCHCYPNVDLV 357
QY 359 DGEVCEVDPDFRANCYQCPNQTSLYLCVCAEGEAPIPHEPRCOMFCNQTACPADCD 418
DB 359 DGEVCEVDPDFRANCYQCPNQTSLYLCVCAEGEAPIPHEPRCOMFCNQTACPADCD 418
QY 358 DGEVCEVDPDFRANCYQCPNQTSLYLCVCAEGEAPIPHEPRCOMFCNQTACPADCD 417
DB 358 DGEVCEVDPDFRANCYQCPNQTSLYLCVCAEGEAPIPHEPRCOMFCNQTACPADCD 417
QY 419 PNTQASCEPEGVLDGFTCTDIDECENGFGCGVCHNLPDGTFCICIGDSDALARRHIGT 478
DB 419 PNTQASCEPEGVLDGFTCTDIDECENGFGCGVCHNLPDGTFCICIGDSDALARRHIGT 478
QY 418 PNSTVCECEPGLDGSVCTDIDECSECEFTSECRNPFPGYECICIGDSDALARRHIGT 477
DB 418 PNSTVCECEPGLDGSVCTDIDECSECEFTSECRNPFPGYECICIGDSDALARRHIGT 477
QY 479 DCDGSKV---DGDGSGSGEPPTPGSTLTTPPAVLVHSG 516
DB 479 DCDGSKV---DGDGSGSGEPPTPGSTLTTPPAVLVHSG 516
QY 478 DCDPIPVREDTKEEGSGEPVPTPGSPTGPPSARPVHSG 518
DB 478 DCDPIPVREDTKEEGSGEPVPTPGSPTGPPSARPVHSG 518

RESULT 3
TRBM_BOVIN
ID TRBM_BOVIN STANDARD; PRT; 356 AA.
AC P06579;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE THROMBOMODULIN (FETOMODULIN) (TM) (FRAGMENT).
GN THBD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87067408; PubMed=3024152;
RA Jackman R.W., Beeler D.L., Vandewater L., Rosenberg R.D.;
RT "Characterization of a thrombomodulin cDNA reveals structural
RT similarity to the low density lipoprotein receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 83:8834-8838(1986).
CC -!- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR
CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS
CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE
CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA
CC SCISSIONS THE ACTIVATED COPACTORS OF THE COAGULATION MECHANISM,
CC FACTOR VA AND FACTOR VIIIa, AND THEREBY REDUCES THE AMOUNT OF
```

QY 349 CHCYPNDLVGECVPEPVDPCFRANCEYQCOPLNQTSLVCAEGFAPIPHPHRCOMFC 408
 ||| | :|||||:||||| |||||: : : :|:|||||:| ||:|||||
 Db 123 CHCOTGVELVDGCVDPVDFDNCCEYQCPVGESEHKCAEGFAPVPAPGPHKQMF 182
 ||| | :|||||:||||| |||||: : : :|:|||||:| ||:|||||
 QY 409 NOTACPADPNTQASCEPCGYILDDGFICTIDCEGNGFCGVCVCHNLPGTFPCICGP 468
 ||| | :|||||:||||| |||||: : : :|:|||||:| ||:|||||
 Db 183 NOTSCPADCPHYPTICRCEGYIIDEGSTCTDINECDT-NICPGCHNLPGTYECICGP 241
 ||| | :|||||:||||| |||||: : : :|:|||||:| ||:|||||
 QY 469 DSALARHGTGDCSGKV-----DGDGSGSGEPPTGPGSTLTP-PA-VGLVHSG 516
 ||||: || ||| :| | ||||| |||||: | || | |||
 Db 242 DSALSGQIGIDCDPTQVNEERGTEPDYG--GSGEPPVPTGATARPSPAPAGPLHSG 297
 ||||: || ||| :| | ||||| |||||: | || | |||

RESULT 4
 FBN2_MOUSE
 ID FBN2_MOUSE STANDARD; PRT; 2907 AA.
 AC Q61555; Q63957;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE FIBRILLIN 2 PRECURSOR.
 GN FBN2 OR FBN-2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95263670; PubMed=7744963;
 RA Zhang H., Hu W., Ramirez F.;
 RT "Developmental expression of fibrillin genes suggests heterogeneity
 of extracellular microfibrils.";
 RL J. Cell Biol. 129:1165-1176(1995).
 RN [2]
 RP SEQUENCE OF 210-317 FROM N.A.
 RX MEDLINE=94140368; PubMed=8307578;
 RA Li X., Pereira L., Zhang H., Sanguinetti C., Ramirez F., Bonadio J.,
 RA Francke U.;
 RT "Fibrillin genes map to regions of conserved mouse/human synteny on
 mouse chromosomes 2 and 18.";
 RL Genomics 18:667-672(1993)
 CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
 THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
 LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
 CC -!- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L39790; AAA74908.1; -;
 CC EMBL; S69359; AAC60685.1; -;
 CC MGD; MGI:95490; Fbn2.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001438; EGF_III.
 DR InterPro; IPR002212; TB.
 DR InterPro; IPR000822; Znf-C2H2.
 DR Pfam; PF00008; EGF; 46.
 DR Pfam; PF00683; TB; 9.
 DR PRINTS; PR00010; EGFLOOD.
 DR SMART; SM00179; EGF_CA; 43.
 DR SMART; SM00001; EGF_like; 3.
 DR PROSITE; PS00010; ASX_HYDROXYL; 43.
 DR PROSITE; PS00022; EGF_1; 2.

DR PROSITE; PS01186; EGF_2; 36.
 DR PROSITE; PS01187; EGF_CA; 43.
 KW Extracellular matrix: Calcium-binding; Glycoprotein; EGF-like domain;
 KW Repeat; Signal; Multigene family.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 2907 FIBRILLIN 2.
 FT DOMAIN 111 142 EGF-LIKE 1, NON-CALCIUM BINDING.
 FT DOMAIN 145 176 EGF-LIKE 2, NON-CALCIUM BINDING.
 FT DOMAIN 176 208 EGF-LIKE 3, NON-CALCIUM BINDING.
 FT DOMAIN 276 317 EGF-LIKE 4, CALCIUM-BINDING.
 FT DOMAIN 318 359 EGF-LIKE 5, CALCIUM-BINDING.
 FT REPEAT 360 426 TGFBP 1.
 FT DOMAIN 487 527 EGF-LIKE 6, NON-CALCIUM BINDING.
 FT DOMAIN 528 567 EGF-LIKE 7, CALCIUM-BINDING.
 FT DOMAIN 568 609 EGF-LIKE 8, CALCIUM-BINDING.
 FT DOMAIN 610 650 EGF-LIKE 9, CALCIUM-BINDING.
 FT DOMAIN 651 691 EGF-LIKE 10, CALCIUM-BINDING.
 FT REPEAT 760 760 TGFBP 2.
 FT DOMAIN 761 802 EGF-LIKE 11, CALCIUM-BINDING.
 FT DOMAIN 803 844 EGF-LIKE 12, CALCIUM-BINDING.
 FT DOMAIN 845 883 EGF-LIKE 13, CALCIUM-BINDING.
 FT DOMAIN 948 989 EGF-LIKE 14, CALCIUM-BINDING.
 FT REPEAT 990 1065 TGFBP 3.
 FT DOMAIN 1066 1107 EGF-LIKE 15, CALCIUM-BINDING.
 FT DOMAIN 1108 1150 EGF-LIKE 16, CALCIUM-BINDING.
 FT DOMAIN 1151 1192 EGF-LIKE 17, CALCIUM-BINDING.
 FT DOMAIN 1193 1234 EGF-LIKE 18, CALCIUM-BINDING.
 FT DOMAIN 1235 1275 EGF-LIKE 19, CALCIUM-BINDING.
 FT DOMAIN 1276 1317 EGF-LIKE 20, CALCIUM-BINDING.
 FT DOMAIN 1318 1359 EGF-LIKE 21, CALCIUM-BINDING.
 FT DOMAIN 1360 1400 EGF-LIKE 22, CALCIUM-BINDING.
 FT DOMAIN 1401 1441 EGF-LIKE 23, CALCIUM-BINDING.
 FT DOMAIN 1442 1483 EGF-LIKE 24, CALCIUM-BINDING.
 FT DOMAIN 1484 1524 EGF-LIKE 25, CALCIUM-BINDING.
 FT DOMAIN 1525 1565 EGF-LIKE 26, CALCIUM-BINDING.
 FT REPEAT 1566 1642 TGFBP 4.
 FT DOMAIN 1643 1684 EGF-LIKE 27, CALCIUM-BINDING.
 FT DOMAIN 1685 1726 EGF-LIKE 28, CALCIUM-BINDING.
 FT REPEAT 1727 1800 TGFBP 5.
 FT DOMAIN 1801 1842 EGF-LIKE 29, CALCIUM-BINDING.
 FT DOMAIN 1843 1884 EGF-LIKE 30, CALCIUM-BINDING.
 FT DOMAIN 1885 1926 EGF-LIKE 31, CALCIUM-BINDING.
 FT DOMAIN 1927 1965 EGF-LIKE 32, CALCIUM-BINDING.
 FT DOMAIN 1966 2008 EGF-LIKE 33, CALCIUM-BINDING.
 FT DOMAIN 2009 2048 EGF-LIKE 34, CALCIUM-BINDING.
 FT DOMAIN 2049 2090 EGF-LIKE 35, CALCIUM-BINDING.
 FT REPEAT 2091 2163 TGFBP 6.
 FT DOMAIN 2164 2205 EGF-LIKE 36, CALCIUM-BINDING.
 FT DOMAIN 2206 2245 EGF-LIKE 37, CALCIUM-BINDING.
 FT DOMAIN 2246 2286 EGF-LIKE 38, CALCIUM-BINDING.
 FT DOMAIN 2287 2330 EGF-LIKE 39, CALCIUM-BINDING.
 FT DOMAIN 2331 2372 EGF-LIKE 40, CALCIUM-BINDING.
 FT REPEAT 2373 2441 TGFBP 7.
 FT DOMAIN 2442 2483 EGF-LIKE 41, CALCIUM-BINDING.
 FT DOMAIN 2484 2524 EGF-LIKE 42, CALCIUM-BINDING.
 FT DOMAIN 2525 2563 EGF-LIKE 43, CALCIUM-BINDING.
 FT DOMAIN 2564 2606 EGF-LIKE 44, CALCIUM-BINDING.
 FT DOMAIN 2607 2646 EGF-LIKE 45, CALCIUM-BINDING.
 FT DOMAIN 2647 2687 EGF-LIKE 46, CALCIUM-BINDING.
 FT DOMAIN 2688 2727 EGF-LIKE 47, CALCIUM-BINDING.
 FT DISULFID 115 124 BY SIMILARITY.
 FT DISULFID 119 130 BY SIMILARITY.
 FT DISULFID 132 141 BY SIMILARITY.
 FT DISULFID 149 159 BY SIMILARITY.
 FT DISULFID 153 164 BY SIMILARITY.
 FT DISULFID 166 175 BY SIMILARITY.
 FT DISULFID 180 190 BY SIMILARITY.
 FT DISULFID 184 196 BY SIMILARITY.
 FT DISULFID 198 207 BY SIMILARITY.
 FT DISULFID 280 292 BY SIMILARITY.
 FT DISULFID 287 301 BY SIMILARITY.
 FT DISULFID 303 316 BY SIMILARITY.
 FT DISULFID 322 334 BY SIMILARITY.

or send an email to license@isb-sib.ch).

CC EMBL; L28748; AAA74122.1; --
DR HSP; P35555; IAPJ
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR002212; TB.
DR Pfam; PF00008; EGF; 46.
DR Pfam; PF00683; TB; 9.
DR PRINTS; PR00010; EGFBL00D.
DR SMART; SM00179; EGF_Ca; 42.
DR SMART; SM00001; EGF_like; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 38.
DR PROSITE; PS01187; EGF_Ca; 45.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 2871 FIBRILLIN 1.
FT DOMAIN 81 112 EGF-LIKE 1, NON-CALCIUM BINDING.
FT DOMAIN 115 146 EGF-LIKE 2, NON-CALCIUM BINDING.
FT DOMAIN 147 178 EGF-LIKE 3, NON-CALCIUM BINDING.
FT DOMAIN 246 287 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 288 329 EGF-LIKE 5, CALCIUM-BINDING.
FT REPEAT 330 390 TGFBB 1.
FT DOMAIN 392 446 PRO-RICH.
FT DOMAIN 449 489 EGF-LIKE 6, NON-CALCIUM BINDING.
FT DOMAIN 490 529 EGF-LIKE 7, CALCIUM-BINDING.
FT DOMAIN 530 571 EGF-LIKE 8, CALCIUM-BINDING.
FT DOMAIN 572 612 EGF-LIKE 9, CALCIUM-BINDING.
FT DOMAIN 613 653 EGF-LIKE 10, CALCIUM-BINDING.
FT REPEAT 654 722 TGFBB 2.
FT DOMAIN 723 764 EGF-LIKE 11, CALCIUM-BINDING.
FT DOMAIN 765 806 EGF-LIKE 12, CALCIUM-BINDING.
FT DOMAIN 807 846 EGF-LIKE 13, CALCIUM-BINDING.
FT DOMAIN 910 951 EGF-LIKE 14, CALCIUM-BINDING.
FT REPEAT 952 1027 TGFBB 3.
FT DOMAIN 1028 1069 EGF-LIKE 15, CALCIUM-BINDING.
FT DOMAIN 1070 1112 EGF-LIKE 16, CALCIUM-BINDING.
FT DOMAIN 1113 1154 EGF-LIKE 17, CALCIUM-BINDING.
FT DOMAIN 1155 1196 EGF-LIKE 18, CALCIUM-BINDING.
FT DOMAIN 1197 1237 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1238 1279 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1280 1321 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1322 1362 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1363 1403 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1404 1445 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1446 1486 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1487 1527 EGF-LIKE 26, CALCIUM-BINDING.
FT REPEAT 1528 1605 TGFBB 4.
FT DOMAIN 1606 1647 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1648 1688 EGF-LIKE 28, CALCIUM-BINDING.
FT REPEAT 1689 1765 TGFBB 5.
FT DOMAIN 1766 1807 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1808 1848 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1849 1890 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1891 1929 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1930 1972 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 1973 2012 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2013 2054 EGF-LIKE 35, CALCIUM-BINDING.
FT REPEAT 2055 2126 TGFBB 6.
FT DOMAIN 2127 2165 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2166 2205 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2206 2246 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2247 2290 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2291 2332 EGF-LIKE 40, CALCIUM-BINDING.
FT REPEAT 2333 2401 TGFBB 7.
FT DOMAIN 2402 2443 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2444 2484 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2485 2523 EGF-LIKE 43, CALCIUM-BINDING.

FT DOMAIN 2524 2566 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2567 2606 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2607 2647 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2648 2687 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFID 85 94 BY SIMILARITY.
FT DISULFID 89 100 BY SIMILARITY.
FT DISULFID 102 111 BY SIMILARITY.
FT DISULFID 119 129 BY SIMILARITY.
FT DISULFID 123 134 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 150 160 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
FT DISULFID 168 177 BY SIMILARITY.
FT DISULFID 250 262 BY SIMILARITY.
FT DISULFID 257 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 304 BY SIMILARITY.
FT DISULFID 299 313 BY SIMILARITY.
FT DISULFID 315 328 BY SIMILARITY.
FT DISULFID 453 465 BY SIMILARITY.
FT DISULFID 460 474 BY SIMILARITY.
FT DISULFID 476 488 BY SIMILARITY.
FT DISULFID 494 504 BY SIMILARITY.
FT DISULFID 499 513 BY SIMILARITY.
FT DISULFID 515 528 BY SIMILARITY.
FT DISULFID 534 546 BY SIMILARITY.
FT DISULFID 541 555 BY SIMILARITY.
FT DISULFID 557 570 BY SIMILARITY.
FT DISULFID 576 587 BY SIMILARITY.
FT DISULFID 582 596 BY SIMILARITY.
FT DISULFID 598 611 BY SIMILARITY.
FT DISULFID 617 628 BY SIMILARITY.
FT DISULFID 623 637 BY SIMILARITY.
FT DISULFID 639 652 BY SIMILARITY.
FT DISULFID 727 739 BY SIMILARITY.
FT DISULFID 734 748 BY SIMILARITY.
FT DISULFID 750 763 BY SIMILARITY.
FT DISULFID 769 781 BY SIMILARITY.
FT DISULFID 776 790 BY SIMILARITY.
FT DISULFID 792 805 BY SIMILARITY.
FT DISULFID 811 821 BY SIMILARITY.
FT DISULFID 816 830 BY SIMILARITY.
FT DISULFID 832 845 BY SIMILARITY.
FT DISULFID 911 926 BY SIMILARITY.
FT DISULFID 921 935 BY SIMILARITY.
FT DISULFID 937 950 BY SIMILARITY.
FT DISULFID 1032 1044 BY SIMILARITY.
FT DISULFID 1039 1053 BY SIMILARITY.
FT DISULFID 1055 1068 BY SIMILARITY.
FT DISULFID 1074 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1111 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1124 1138 BY SIMILARITY.
FT DISULFID 1140 1153 BY SIMILARITY.
FT DISULFID 1159 1171 BY SIMILARITY.
FT DISULFID 1166 1180 BY SIMILARITY.
FT DISULFID 1182 1195 BY SIMILARITY.
FT DISULFID 1201 1212 BY SIMILARITY.
FT DISULFID 1208 1221 BY SIMILARITY.
FT DISULFID 1223 1236 BY SIMILARITY.
FT DISULFID 1242 1254 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1265 1278 BY SIMILARITY.
FT DISULFID 1284 1296 BY SIMILARITY.
FT DISULFID 1291 1305 BY SIMILARITY.
FT DISULFID 1307 1320 BY SIMILARITY.
FT DISULFID 1326 1339 BY SIMILARITY.
FT DISULFID 1333 1348 BY SIMILARITY.
FT DISULFID 1350 1361 BY SIMILARITY.
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.
FT DISULFID 1391 1402 BY SIMILARITY.


```
FT DISULFID 928 941 BY SIMILARITY.
FT DISULFID 947 956 BY SIMILARITY.
FT DISULFID 952 965 BY SIMILARITY.
FT DISULFID 967 980 BY SIMILARITY.
FT DISULFID 986 998 BY SIMILARITY.
FT DISULFID 994 1007 BY SIMILARITY.
FT DISULFID 1009 1023 BY SIMILARITY.
FT DISULFID 1029 1042 BY SIMILARITY.
FT DISULFID 1036 1051 BY SIMILARITY.
FT DISULFID 1056 1068 BY SIMILARITY.
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1035 1035 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1184 AA; 126543 MW; CA48490A55F9EC5D CRC64;

Query Match 12.6%; Score 367.5; DB 1; Length 1184;
Best Local Similarity 30.0%; Pred. No. 3.8e-17;
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;

QY 154 EQOCEVKAD-GFCEPHEFPATC-----RPLAV-----EPGAAAAVSITYGTPFAARG 201
Db 531 EQSCESNPNGYPCN-HVMLSCEGEPLTVPEVRRPPEAAAPRRVS-----EAEM 582
QY 202 ADFCALPVGSSAAV---APLGLQLMCTAPPAGVGHWAAREAPGAWDCSVENGCEHACNA 258
Db 583 AGRALSISGTEAEPLNSLPGDDQDECLLPGEL-----COHLGIN 622
QY 259 IPGAPRCQCPAGAAALQADGRSC-----TASAT-----QS 287
Db 623 TVGSYHCACFPFSLQDDGRTCRPEGHPPQPEAPQPAKSEFSQVANSNTIPLPQPN 682
QY 288 CND--LCEHEFCVPNDQFQSVCMCTGYRLAADOHRCEVDVDDCILEPSPCP--QRCVNT 343
Db 683 CKDNGPKQVC--STVGSSAICSCFPGYALMADGVSCDINECVTLHTCSRGEHCVNT 739
QY 344 QGFECH----CYPNYDLVDCVEPDPDPCFRANCEYQCQPLNQTSLYCV----- 389
Db 740 LGSFHCYKALTCEPGYALKDGECE-EDVDEC--AMGTHTCQP-----GFLCQNTKGSFYCQA 792
QY 390 ---CAEGFAPIPH-----EPHRCMFNCNOTACPADCDPNTPQASCECPGV-I 432
Db 793 RQRCMDGFLQDPGNCVDINECTSLSEPCRGFSCINTVGSYTCORNPILI---CARGIHA 849
QY 433 LDDGFICTIDECENGGFCSG---VCHNLPGTFECIC 466
Db 850 SDDGAKVDVNECETGVHRCGEGQVCHNLPGSYRDC 886

RESULT 7
FBNI_PIG
ID FBNI_PIG STANDARD; PRT; 2871 AA.
AC Q9TV36;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE FBIRILLIN 1 PRECURSOR.
GN FBNI.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=99156858; PubMed=10036187;
RA Bieri N.J., Eldadah Z.A., Moore C.S., Stetten G., Spencer F.,
RA Dietz H.C.;
RT "Revised genomic organization of FBNI and significance for regulated
RT gene expression.";
RL Genomics 56:70-77(1999).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC -!- FUNCTION: BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
```

```
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -!- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC MICROFIBRILS.
CC -!- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF073800; AAD50328.1; -.
CC HSSP; P35555; IAPJ.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR001438; EGF-II.
DR InterPro; IPR002212; TB.
DR InterPro; IPR000822; Znf-C2H2.
DR Pfam; PF00008; EGF; 45.
DR Pfam; PF00683; TB; 9.
DR PRINTS; PR00010; EGFBL00D.
DR SMART; SM00179; EGF_CA; 40.
DR SMART; SM00001; EGF_Like; 6.
DR PROSITE; PS00010; ASX_HYDROXYL; 41.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 36.
DR PROSITE; PS01187; EGF_CA; 41.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 2871 FIBRILLIN 1.
FT DOMAIN 81 112 EGF-LIKE 1, NON-CALCIUM BINDING.
FT DOMAIN 115 146 EGF-LIKE 2, NON-CALCIUM BINDING.
FT DOMAIN 147 178 EGF-LIKE 3, NON-CALCIUM BINDING.
FT DOMAIN 246 287 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 288 329 EGF-LIKE 5, CALCIUM-BINDING.
FT REPEAT 330 390 TGFBP 1.
FT DOMAIN 392 446 PRO-RICH.
FT DOMAIN 449 489 EGF-LIKE 6, NON-CALCIUM BINDING.
FT DOMAIN 490 529 EGF-LIKE 7, CALCIUM-BINDING.
FT DOMAIN 530 571 EGF-LIKE 8, CALCIUM-BINDING.
FT DOMAIN 572 612 EGF-LIKE 9, CALCIUM-BINDING.
FT DOMAIN 613 653 EGF-LIKE 10, CALCIUM-BINDING.
FT REPEAT 654 722 TGFBP 2.
FT DOMAIN 723 764 EGF-LIKE 11, CALCIUM-BINDING.
FT DOMAIN 765 806 EGF-LIKE 12, CALCIUM-BINDING.
FT DOMAIN 807 846 EGF-LIKE 13, CALCIUM-BINDING.
FT DOMAIN 910 951 EGF-LIKE 14, CALCIUM-BINDING.
FT REPEAT 952 1027 TGFBP 3.
FT DOMAIN 1028 1069 EGF-LIKE 15, CALCIUM-BINDING.
FT DOMAIN 1070 1112 EGF-LIKE 16, CALCIUM-BINDING.
FT DOMAIN 1113 1154 EGF-LIKE 17, CALCIUM-BINDING.
FT DOMAIN 1155 1196 EGF-LIKE 18, CALCIUM-BINDING.
FT DOMAIN 1197 1237 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1238 1279 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1280 1321 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1322 1362 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1363 1403 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1404 1445 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1446 1487 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1487 1527 EGF-LIKE 26, CALCIUM-BINDING.
FT REPEAT 1528 1605 TGFBP 4.
FT DOMAIN 1606 1647 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1648 1688 EGF-LIKE 28, CALCIUM-BINDING.
FT REPEAT 1689 1765 TGFBP 5.
FT DOMAIN 1766 1807 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1808 1848 EGF-LIKE 30, CALCIUM-BINDING.
```

```
FT DOMAIN 1849 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1891 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1930 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 1972 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2012 EGF-LIKE 35, CALCIUM-BINDING.
FT DOMAIN 2054 TGRBP 6.
FT REPEAT 2055 EGF-LIKE 36, CALCIUM-BINDING.
FT REPEAT 2126 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2166 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2206 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2247 EGF-LIKE 40, CALCIUM-BINDING.
FT REPEAT 2291 TGRBP 7.
FT REPEAT 2332 EGF-LIKE 41, CALCIUM-BINDING.
FT REPEAT 2333 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2402 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2444 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2485 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2524 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2567 EGF-LIKE 47, CALCIUM-BINDING.
FT DOMAIN 2607
FT DOMAIN 2648
FT DISULFID 85
FT DISULFID 89
FT DISULFID 102
FT DISULFID 102
FT DISULFID 119
FT DISULFID 123
FT DISULFID 136
FT DISULFID 150
FT DISULFID 154
FT DISULFID 168
FT DISULFID 177
FT DISULFID 250
FT DISULFID 257
FT DISULFID 273
FT DISULFID 292
FT DISULFID 299
FT DISULFID 313
FT DISULFID 315
FT DISULFID 453
FT DISULFID 460
FT DISULFID 476
FT DISULFID 494
FT DISULFID 499
FT DISULFID 515
FT DISULFID 534
FT DISULFID 541
FT DISULFID 555
FT DISULFID 557
FT DISULFID 576
FT DISULFID 582
FT DISULFID 582
FT DISULFID 598
FT DISULFID 611
FT DISULFID 617
FT DISULFID 623
FT DISULFID 639
FT DISULFID 727
FT DISULFID 734
FT DISULFID 750
FT DISULFID 769
FT DISULFID 776
FT DISULFID 792
FT DISULFID 811
FT DISULFID 821
FT DISULFID 830
FT DISULFID 832
FT DISULFID 914
FT DISULFID 921
FT DISULFID 935
FT DISULFID 937
FT DISULFID 1032
FT DISULFID 1039
FT DISULFID 1053
FT DISULFID 1068
FT DISULFID 1074
FT DISULFID 1081
FT DISULFID 1095
FT DISULFID 1097
FT DISULFID 1117
FT DISULFID 1124
FT DISULFID 1138
FT DISULFID 1140
FT DISULFID 1159
FT DISULFID 1171
FT DISULFID 1201
FT DISULFID 1208
FT DISULFID 1212
FT DISULFID 1221
```

```
FT DISULFID 1223 1236 BY SIMILARITY.
FT DISULFID 1242 1254 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1265 1278 BY SIMILARITY.
FT DISULFID 1284 1296 BY SIMILARITY.
FT DISULFID 1291 1305 BY SIMILARITY.
FT DISULFID 1307 1320 BY SIMILARITY.
FT DISULFID 1326 1339 BY SIMILARITY.
FT DISULFID 1333 1348 BY SIMILARITY.
FT DISULFID 1350 1361 BY SIMILARITY.
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.
FT DISULFID 1391 1402 BY SIMILARITY.
FT DISULFID 1408 1420 BY SIMILARITY.
FT DISULFID 1415 1429 BY SIMILARITY.
FT DISULFID 1450 1461 BY SIMILARITY.
FT DISULFID 1456 1470 BY SIMILARITY.
FT DISULFID 1472 1485 BY SIMILARITY.
FT DISULFID 1491 1502 BY SIMILARITY.
FT DISULFID 1497 1511 BY SIMILARITY.
FT DISULFID 1513 1526 BY SIMILARITY.
FT DISULFID 1610 1622 BY SIMILARITY.
FT DISULFID 1617 1631 BY SIMILARITY.
FT DISULFID 1633 1646 BY SIMILARITY.
FT DISULFID 1652 1663 BY SIMILARITY.
FT DISULFID 1658 1672 BY SIMILARITY.
FT DISULFID 1674 1687 BY SIMILARITY.
FT DISULFID 1770 1782 BY SIMILARITY.
FT DISULFID 1777 1791 BY SIMILARITY.
FT DISULFID 1793 1806 BY SIMILARITY.
FT DISULFID 1812 1824 BY SIMILARITY.
FT DISULFID 1818 1833 BY SIMILARITY.

Query Match 12.6%; Score 366; DB 1; Length 2871;
Best Local Similarity 30.6%; Pred. No. 1.le-16;
Matches 109; Conservative 26; Mismatches 107; Indels 114; Gaps 20;

QY 244 DCSVENGGCEHACNAIPGAPRCQCPAGAAQAQADORSCTASATQSCNDLCEHFCVNP--- 300
DB 1200 ECSIMNGGCTFTCTNSEGSEYECSCQPGFALPMDQRST-----DIDE--CEDNPNIC 1249

QY 301 -----DQPGSYSCMCETGYRLAADQHRCEDDVDCILEPSPC-PQRCVNTGGFECHCYP 353
DB 1250 DGGCTNIPGEYRCLCYDGFMASEDMTCTVDVNECDLNPICLSGTCTENTKGSFICHCDM 1309

QY 354 NYDLVDGE--CVEPVDP--FRANCEYQCOPLNOT--SYLCVCARGFA-----PIP 398
DB 1310 GYSGKKGKTGCTD-INECEIGAHNCDRHAVCTNTAGSFNCSCSPGWIGDGIKCTDLDECS 1368

QY 399 HEPHRCQFNCQTACPADCDPNTQAS--CECPGYILDDGFICTDIDEC--EN----- 447
DB 1369 NGTHMCSQH-----ADC-KNTMGSYRCLCKEY--TGDGFTCADLDECSENVKLCGNVQ 1419

QY 448 -----GGF-----CS-----GVCHNLPGTFFECICGPDALA 473
DB 1420 CLYAPGGYHCEYDMGFVPSADRKSCVDSDECSLNPICVFTGCHNLPGLFRCEC----- 1472

QY 474 RHIGTDCDCKVDGDSGSGEPPPS-----PTGPS-----TLTPPAVGLV 513
DB 1473 -EIGYELDRSGGCTDVNECLEPPTCISGNCVNTPGSYTCVCPDPFELNTRVGCV 1527
```

RESULT 8

```
FBN2_HUMAN
ID FBN2_HUMAN STANDARD; PRT; 2911 AA.
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE FIBRILLIN 2 PRECURSOR.
GN FBN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```


OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94165150; PubMed=8120105;
 RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sanguineti C.,
 RA Bonadio J., Mecham R.P., Ramirez F.,
 RT "Structure and expression of fibrillin-2, a novel microfibrillar
 RT component preferentially located in elastic matrices";
 RL J. Cell Biol. 124:855-863(1994).
 RN [2]
 RN SEQUENCE OF 752-1505 FROM N.A.
 RP MEDLINE=91304567; PubMed=1852206;
 RX Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,
 RA Tsipouras P., Ramirez F., Hollister D.;
 RT "Linkage of Marfan syndrome and a phenotypically related disorder to
 RT two different fibrillin genes.";
 RL Nature 352:330-334(1991).
 RN [3]
 RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.
 RX MEDLINE=96083599; PubMed=7493032;
 RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;
 RT "Fibrillin-2 (FN2) mutations result in the Marfan-like disorder,
 RT congenital contractural arachnoidactyly.";
 RL Nat. Genet. 11:456-458(1995).
 RN [4]
 RP VARIANTS CCA HIS-1114.
 RX MEDLINE=98407789; PubMed=9737771;
 RA Babcock D., Gasner C., Francke U., Maslen C.;
 RT "A single mutation that results in an asp-to-his substitution and
 RT partial exon skipping in a family with congenital contractural
 RT arachnoidactyly.";
 RL Hum. Genet. 103:22-28(1998).
 RN [5]
 RP VARIANTS CCA PHE-1141 AND TRP-1252.
 RX MEDLINE=20259236; PubMed=10797416;
 RA Belleh S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,
 RA Godfrey M.;
 RT "Two novel fibrillin-2 mutations in congenital contractural
 RT arachnoidactyly.";
 RL Am. J. Med. Genet. 92:7-12(2000).
 CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
 CC THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE
 CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.
 CC -1- DISEASE: DEFECTS IN FN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL
 CC ARACHNOIDACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS
 CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE
 CC AORTA AND THE EYES.
 CC -1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
 CC EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U03272; AAA18950.1; -
 DR EMBL; X62009; -; NOT_ANNOTATED_CDS.
 DR PIR; S17063; S17063.
 DR PIR; S31101; S31101.
 DR HSP; P35555; IEMO.
 DR MIM; 121050; -
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001438; EGF_II.
 DR InterPro; IPR002212; TB.
 DR Pfam; PF00008; EGF; 46.
 DR Pfam; PF00683; TB; 9.

DR PRINTS; PRO0010; EGFBL00D.
 DR SMART; SM00179; EGF_CA; 43.
 DR SMART; SM00001; EGF_like; 3.
 DR PROSITE; PS00010; ASX_HYDROXYL; 43.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 37.
 DR PROSITE; PS01187; EGF_CA; 43.
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
 KW Repeat; Signal; Multigene family; Disease mutation; Polymorphism.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 2911 FIBRILLIN 2.
 FT DOMAIN 111 142 EGF-LIKE 1, NON-CALCIUM BINDING.
 FT DOMAIN 145 176 EGF-LIKE 2, NON-CALCIUM BINDING.
 FT DOMAIN 176 207 EGF-LIKE 3, NON-CALCIUM BINDING.
 FT DOMAIN 275 316 EGF-LIKE 4, CALCIUM-BINDING.
 FT DOMAIN 317 358 EGF-LIKE 5, CALCIUM-BINDING.
 FT REPEAT 359 425 TGFBP 1.
 FT DOMAIN 493 533 EGF-LIKE 6, NON-CALCIUM BINDING.
 FT DOMAIN 534 573 EGF-LIKE 7, CALCIUM-BINDING.
 FT DOMAIN 574 615 EGF-LIKE 8, CALCIUM-BINDING.
 FT DOMAIN 616 656 EGF-LIKE 9, CALCIUM-BINDING.
 FT DOMAIN 657 697 EGF-LIKE 10, CALCIUM-BINDING.
 FT REPEAT 698 766 TGFBP 2.
 FT DOMAIN 767 808 EGF-LIKE 11, CALCIUM-BINDING.
 FT DOMAIN 809 850 EGF-LIKE 12, CALCIUM-BINDING.
 FT DOMAIN 851 890 EGF-LIKE 13, CALCIUM-BINDING.
 FT DOMAIN 894 995 EGF-LIKE 14, CALCIUM-BINDING.
 FT REPEAT 996 1071 TGFBP 3.
 FT DOMAIN 1072 1113 EGF-LIKE 15, CALCIUM-BINDING.
 FT DOMAIN 1114 1156 EGF-LIKE 16, CALCIUM-BINDING.
 FT DOMAIN 1157 1198 EGF-LIKE 17, CALCIUM-BINDING.
 FT DOMAIN 1199 1240 EGF-LIKE 18, CALCIUM-BINDING.
 FT DOMAIN 1241 1281 EGF-LIKE 19, CALCIUM-BINDING.
 FT DOMAIN 1282 1323 EGF-LIKE 20, CALCIUM-BINDING.
 FT DOMAIN 1324 1365 EGF-LIKE 21, CALCIUM-BINDING.
 FT DOMAIN 1366 1406 EGF-LIKE 22, CALCIUM-BINDING.
 FT DOMAIN 1407 1447 EGF-LIKE 23, CALCIUM-BINDING.
 FT DOMAIN 1448 1489 EGF-LIKE 24, CALCIUM-BINDING.
 FT DOMAIN 1490 1530 EGF-LIKE 25, CALCIUM-BINDING.
 FT DOMAIN 1531 1571 EGF-LIKE 26, CALCIUM-BINDING.
 FT REPEAT 1572 1648 TGFBP 4.
 FT DOMAIN 1649 1690 EGF-LIKE 27, CALCIUM-BINDING.
 FT DOMAIN 1691 1732 EGF-LIKE 28, CALCIUM-BINDING.
 FT REPEAT 1733 1806 TGFBP 5.
 FT DOMAIN 1807 1848 EGF-LIKE 29, CALCIUM-BINDING.
 FT DOMAIN 1849 1890 EGF-LIKE 30, CALCIUM-BINDING.
 FT DOMAIN 1891 1932 EGF-LIKE 31, CALCIUM-BINDING.
 FT DOMAIN 1933 1971 EGF-LIKE 32, CALCIUM-BINDING.
 FT DOMAIN 1972 2014 EGF-LIKE 33, CALCIUM-BINDING.
 FT DOMAIN 2015 2054 EGF-LIKE 34, CALCIUM-BINDING.
 FT DOMAIN 2055 2096 EGF-LIKE 35, CALCIUM-BINDING.
 FT REPEAT 2097 2169 TGFBP 6.
 FT DOMAIN 2170 2211 EGF-LIKE 36, CALCIUM-BINDING.
 FT DOMAIN 2212 2251 EGF-LIKE 37, CALCIUM-BINDING.
 FT DOMAIN 2252 2292 EGF-LIKE 38, CALCIUM-BINDING.
 FT DOMAIN 2293 2336 EGF-LIKE 39, CALCIUM-BINDING.
 FT DOMAIN 2337 2378 EGF-LIKE 40, CALCIUM-BINDING.
 FT REPEAT 2379 2447 TGFBP 7.
 FT DOMAIN 2448 2489 EGF-LIKE 41, CALCIUM-BINDING.
 FT DOMAIN 2490 2530 EGF-LIKE 42, CALCIUM-BINDING.
 FT DOMAIN 2531 2569 EGF-LIKE 43, CALCIUM-BINDING.
 FT DOMAIN 2570 2612 EGF-LIKE 44, CALCIUM-BINDING.
 FT DOMAIN 2613 2652 EGF-LIKE 45, CALCIUM-BINDING.
 FT DOMAIN 2653 2693 EGF-LIKE 46, CALCIUM-BINDING.
 FT DOMAIN 2694 2733 EGF-LIKE 47, CALCIUM-BINDING.
 FT DISULFID 115 124 BY SIMILARITY.
 FT DISULFID 119 130 BY SIMILARITY.
 FT DISULFID 132 141 BY SIMILARITY.
 FT DISULFID 149 159 BY SIMILARITY.
 FT DISULFID 153 164 BY SIMILARITY.
 FT DISULFID 166 175 BY SIMILARITY.
 FT DISULFID 180 190 BY SIMILARITY.
 FT DISULFID 184 195 BY SIMILARITY.

| | | | | |
|--|----------|---|------|----------------|
| FT | DISULFID | 197 | 206 | BY SIMILARITY. |
| FT | DISULFID | 279 | 291 | BY SIMILARITY. |
| FT | DISULFID | 286 | 300 | BY SIMILARITY. |
| FT | DISULFID | 302 | 315 | BY SIMILARITY. |
| FT | DISULFID | 321 | 333 | BY SIMILARITY. |
| FT | DISULFID | 328 | 342 | BY SIMILARITY. |
| FT | DISULFID | 344 | 357 | BY SIMILARITY. |
| FT | DISULFID | 497 | 509 | BY SIMILARITY. |
| FT | DISULFID | 504 | 518 | BY SIMILARITY. |
| FT | DISULFID | 520 | 532 | BY SIMILARITY. |
| FT | DISULFID | 538 | 548 | BY SIMILARITY. |
| FT | DISULFID | 543 | 557 | BY SIMILARITY. |
| FT | DISULFID | 559 | 572 | BY SIMILARITY. |
| FT | DISULFID | 578 | 590 | BY SIMILARITY. |
| FT | DISULFID | 585 | 599 | BY SIMILARITY. |
| FT | DISULFID | 601 | 614 | BY SIMILARITY. |
| FT | DISULFID | 620 | 631 | BY SIMILARITY. |
| FT | DISULFID | 626 | 640 | BY SIMILARITY. |
| FT | DISULFID | 642 | 655 | BY SIMILARITY. |
| FT | DISULFID | 661 | 672 | BY SIMILARITY. |
| FT | DISULFID | 667 | 681 | BY SIMILARITY. |
| FT | DISULFID | 683 | 696 | BY SIMILARITY. |
| FT | DISULFID | 771 | 783 | BY SIMILARITY. |
| FT | DISULFID | 778 | 792 | BY SIMILARITY. |
| FT | DISULFID | 794 | 807 | BY SIMILARITY. |
| FT | DISULFID | 813 | 825 | BY SIMILARITY. |
| FT | DISULFID | 820 | 834 | BY SIMILARITY. |
| FT | DISULFID | 836 | 849 | BY SIMILARITY. |
| FT | DISULFID | 855 | 865 | BY SIMILARITY. |
| FT | DISULFID | 860 | 874 | BY SIMILARITY. |
| FT | DISULFID | 876 | 889 | BY SIMILARITY. |
| FT | DISULFID | 958 | 970 | BY SIMILARITY. |
| FT | DISULFID | 965 | 979 | BY SIMILARITY. |
| FT | DISULFID | 981 | 994 | BY SIMILARITY. |
| FT | DISULFID | 1076 | 1088 | BY SIMILARITY. |
| FT | DISULFID | 1083 | 1097 | BY SIMILARITY. |
| FT | DISULFID | 1099 | 1112 | BY SIMILARITY. |
| FT | DISULFID | 1118 | 1130 | BY SIMILARITY. |
| FT | DISULFID | 1125 | 1139 | BY SIMILARITY. |
| FT | DISULFID | 1141 | 1155 | BY SIMILARITY. |
| FT | DISULFID | 1161 | 1173 | BY SIMILARITY. |
| FT | DISULFID | 1168 | 1182 | BY SIMILARITY. |
| FT | DISULFID | 1184 | 1197 | BY SIMILARITY. |
| FT | DISULFID | 1203 | 1215 | BY SIMILARITY. |
| FT | DISULFID | 1210 | 1224 | BY SIMILARITY. |
| Query Match 12.6%; Score 366; DB 1; Length 2911; | | | | |
| Best Local Similarity 33.1%; Pred. No. 1.le-16; | | | | |
| Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps | | | | |
| Qy | 244 | DCSVENGCGCBACNATIPGAPRCQCAGAAALADGRCSCTASATQSCNDLCEHFCVNPDP-- | 301 | |
| Db | 1244 | ECMIMNGGCDTQCTNSEGYSCESEGVAMPDGRSCA-----DIDE--CENNPDIC | 1293 | |
| Qy | 302 | -----QPGYSKCMETGYRLAABDHRCEDVDDCILEPSPCP--QRCVNTQGGFECHCYP | 353 | |
| Db | 1294 | DGGQCTNIPGEYRCLCYDGFMA5MDMKTCIDVNECDLNSNICMFGECENTKGSFICHQCL | 1353 | |
| Qy | 354 | NYDLVJGE--CVPVPDPC--FRANCEYQOQPLN-OTSYLVCVACGFAPIPH-----BPHR | 403 | |
| Db | 1354 | GYSYKGGTTGCTD-VDECEBIGAHNDMHASCLINIPGSKSCREGW--IGNGIKCIDLDE | 1410 | |
| Qy | 404 | QWFCFNQACPADCDNTQAS--CECPGYILDDGFICTDIDE-----CENG----- | 448 | |
| Db | 1411 | CSNGTHQCSINAQC-VNTPGSRKACSEGF-TGDGFTCSYDVECAENINLCENGQCLNVP | 1468 | |
| Qy | 449 | -----GF----- | 477 | |
| Db | 1469 | GAYRCECEMGFTPASDRSCQDIDEC5FQNICVSTCNLFGMFHCICDDGYELDRITGNN | 1528 | |
| Qy | 478 | -TDCD | 481 | |
| Db | 1529 | CTDID | 1533 | |

| | | | |
|----|----------|---|------|
| DR | | PROSITE; PS01186; EGF_2; 38. | |
| DR | | Extracell; PS01187; EGF_CA; 45. | |
| KW | | Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain; | |
| KW | | Repeat; Signal; Multigene family. | |
| FT | SIGNAL | 1 | 27 |
| FT | CHAIN | 28 | 2871 |
| FT | DOMAIN | 81 | 112 |
| FT | DOMAIN | 115 | 146 |
| FT | DOMAIN | 147 | 178 |
| FT | DOMAIN | 246 | 287 |
| FT | DOMAIN | 288 | 329 |
| FT | DOMAIN | 330 | 401 |
| FT | DOMAIN | 402 | 445 |
| FT | DOMAIN | 449 | 489 |
| FT | DOMAIN | 490 | 529 |
| FT | DOMAIN | 530 | 571 |
| FT | DOMAIN | 572 | 612 |
| FT | DOMAIN | 613 | 653 |
| FT | DOMAIN | 656 | 721 |
| FT | DOMAIN | 723 | 764 |
| FT | DOMAIN | 765 | 806 |
| FT | DOMAIN | 807 | 846 |
| FT | DOMAIN | 910 | 951 |
| FT | DOMAIN | 952 | 1018 |
| FT | DOMAIN | 1028 | 1069 |
| FT | DOMAIN | 1070 | 1112 |
| FT | DOMAIN | 1113 | 1154 |
| FT | DOMAIN | 1155 | 1196 |
| ET | DOMAIN | 1197 | 1237 |
| FT | DOMAIN | 1238 | 1279 |
| FT | DOMAIN | 1280 | 1321 |
| FT | DOMAIN | 1322 | 1362 |
| FT | DOMAIN | 1363 | 1403 |
| FT | DOMAIN | 1404 | 1445 |
| FT | DOMAIN | 1446 | 1486 |
| FT | DOMAIN | 1487 | 1527 |
| FT | DOMAIN | 1528 | 1599 |
| FT | DOMAIN | 1606 | 1647 |
| FT | DOMAIN | 1648 | 1688 |
| FT | DOMAIN | 1689 | 1758 |
| FT | DOMAIN | 1766 | 1807 |
| FT | DOMAIN | 1808 | 1848 |
| FT | DOMAIN | 1849 | 1890 |
| FT | DOMAIN | 1891 | 1929 |
| FT | DOMAIN | 1930 | 1972 |
| FT | DOMAIN | 1973 | 2012 |
| FT | DOMAIN | 2013 | 2054 |
| FT | DOMAIN | 2055 | 2121 |
| FT | DOMAIN | 2127 | 2165 |
| FT | DOMAIN | 2166 | 2205 |
| FT | DOMAIN | 2206 | 2246 |
| FT | DOMAIN | 2247 | 2290 |
| FT | DOMAIN | 2291 | 2332 |
| FT | DOMAIN | 2333 | 2400 |
| FT | DOMAIN | 2402 | 2443 |
| FT | DOMAIN | 2444 | 2484 |
| FT | DOMAIN | 2485 | 2523 |
| FT | DOMAIN | 2524 | 2566 |
| FT | DOMAIN | 2567 | 2605 |
| FT | DOMAIN | 2607 | 2647 |
| FT | DOMAIN | 2648 | 2687 |
| FT | DISULFID | 85 | 94 |
| FT | DISULFID | 89 | 100 |
| FT | DISULFID | 102 | 111 |
| FT | DISULFID | 119 | 129 |
| FT | DISULFID | 123 | 134 |
| FT | DISULFID | 136 | 145 |
| FT | DISULFID | 150 | 160 |
| FT | DISULFID | 154 | 166 |
| FT | DISULFID | 168 | 177 |
| FT | DISULFID | 250 | 262 |
| FT | DISULFID | 257 | 271 |
| FT | DISULFID | 273 | 286 |

| | | | | |
|-------------|----------|------|------|--------------------------------------|
| FT | DISULFID | 292 | 304 | BY SIMILARITY. |
| FT | DISULFID | 299 | 313 | BY SIMILARITY. |
| FT | DISULFID | 315 | 328 | BY SIMILARITY. |
| FT | DISULFID | 453 | 465 | BY SIMILARITY. |
| FT | DISULFID | 460 | 474 | BY SIMILARITY. |
| FT | DISULFID | 476 | 488 | BY SIMILARITY. |
| FT | DISULFID | 494 | 504 | BY SIMILARITY. |
| FT | DISULFID | 499 | 513 | BY SIMILARITY. |
| FT | DISULFID | 515 | 528 | BY SIMILARITY. |
| FT | DISULFID | 534 | 546 | BY SIMILARITY. |
| FT | DISULFID | 541 | 555 | BY SIMILARITY. |
| FT | DISULFID | 557 | 570 | BY SIMILARITY. |
| FT | DISULFID | 576 | 587 | BY SIMILARITY. |
| FT | DISULFID | 582 | 596 | BY SIMILARITY. |
| FT | DISULFID | 598 | 611 | BY SIMILARITY. |
| FT | DISULFID | 617 | 628 | BY SIMILARITY. |
| FT | DISULFID | 623 | 637 | BY SIMILARITY. |
| FT | DISULFID | 639 | 652 | BY SIMILARITY. |
| FT | DISULFID | 727 | 739 | BY SIMILARITY. |
| FT | DISULFID | 734 | 748 | BY SIMILARITY. |
| FT | DISULFID | 750 | 763 | BY SIMILARITY. |
| FT | DISULFID | 769 | 781 | BY SIMILARITY. |
| FT | DISULFID | 776 | 790 | BY SIMILARITY. |
| FT | DISULFID | 792 | 805 | BY SIMILARITY. |
| FT | DISULFID | 811 | 821 | BY SIMILARITY. |
| FT | DISULFID | 816 | 830 | BY SIMILARITY. |
| FT | DISULFID | 832 | 845 | BY SIMILARITY. |
| FT | DISULFID | 914 | 926 | BY SIMILARITY. |
| FT | DISULFID | 921 | 935 | BY SIMILARITY. |
| FT | DISULFID | 937 | 950 | BY SIMILARITY. |
| FT | DISULFID | 1032 | 1044 | BY SIMILARITY. |
| FT | DISULFID | 1039 | 1053 | BY SIMILARITY. |
| FT | DISULFID | 1055 | 1068 | BY SIMILARITY. |
| FT | DISULFID | 1074 | 1086 | BY SIMILARITY. |
| FT | DISULFID | 1081 | 1095 | BY SIMILARITY. |
| FT | DISULFID | 1097 | 1111 | BY SIMILARITY. |
| FT | DISULFID | 1117 | 1129 | BY SIMILARITY. |
| FT | DISULFID | 1124 | 1138 | BY SIMILARITY. |
| FT | DISULFID | 1140 | 1153 | BY SIMILARITY. |
| FT | DISULFID | 1159 | 1171 | BY SIMILARITY. |
| FT | DISULFID | 1166 | 1180 | BY SIMILARITY. |
| FT | DISULFID | 1182 | 1195 | BY SIMILARITY. |
| FT | DISULFID | 1201 | 1212 | BY SIMILARITY. |
| FT | DISULFID | 1208 | 1221 | BY SIMILARITY. |
| FT | DISULFID | 1223 | 1236 | BY SIMILARITY. |
| FT | DISULFID | 1242 | 1254 | BY SIMILARITY. |
| FT | DISULFID | 1249 | 1263 | BY SIMILARITY. |
| FT | DISULFID | 1265 | 1278 | BY SIMILARITY. |
| FT | DISULFID | 1284 | 1296 | BY SIMILARITY. |
| FT | DISULFID | 1291 | 1305 | BY SIMILARITY. |
| FT | DISULFID | 1307 | 1320 | BY SIMILARITY. |
| FT | DISULFID | 1326 | 1339 | BY SIMILARITY. |
| FT | DISULFID | 1333 | 1348 | BY SIMILARITY. |
| FT | DISULFID | 1350 | 1361 | BY SIMILARITY. |
| FT | DISULFID | 1367 | 1380 | BY SIMILARITY. |
| FT | DISULFID | 1374 | 1389 | BY SIMILARITY. |
| FT | DISULFID | 1391 | 1402 | BY SIMILARITY. |
| FT | DISULFID | 1408 | 1420 | BY SIMILARITY. |
| FT | DISULFID | 1415 | 1429 | BY SIMILARITY. |
| FT | DISULFID | 1431 | 1444 | BY SIMILARITY. |
| FT | DISULFID | 1450 | 1461 | BY SIMILARITY. |
| FT | DISULFID | 1456 | 1470 | BY SIMILARITY. |
| FT | DISULFID | 1472 | 1485 | BY SIMILARITY. |
| FT | DISULFID | 1491 | 1502 | BY SIMILARITY. |
| FT | DISULFID | 1497 | 1511 | BY SIMILARITY. |
| FT | DISULFID | 1513 | 1526 | BY SIMILARITY. |
| FT | DISULFID | 1610 | 1622 | BY SIMILARITY. |
| FT | DISULFID | 1617 | 1631 | BY SIMILARITY. |
| FT | DISULFID | 1633 | 1646 | BY SIMILARITY. |
| FT | DISULFID | 1652 | 1663 | BY SIMILARITY. |
| FT | DISULFID | 1658 | 1672 | BY SIMILARITY. |
| Query Match | | | | 12.4%; Score 362; DB 1; Length 2871; |

| | | | |
|--|--|--|------|
| Best Local Similarity 29.4%; Pred. No. 1.9e-16; | | | |
| Matches 117; Conservative 26; Mismatches 127; Indels 128; Gaps 21; | | | |
| QY | 212 | SAAVAPLGL-----QLMCTAPPVAGVQHWAREAPGAWDCSVENGSGCEHACNAIPGAPRC | 265 |
| DB | 1162 | SANLCPHGRCVNLKGYQCACNPGYHPTDRLFCVDIDECSIMNGSCETFTCTNSDGSYEC | 1221 |
| QY | 266 | QCPAGAAALQADORSCTASTQSCNDLCEHFCVPNP-----DOPGYSYSCMETGYRL | 316 |
| DB | 1222 | SCQPGFALMPDQRSCT-----DIDQ--CEDPNICDGGQCTNIPGEYRCLCYDGFMA | 1271 |
| QY | 317 | AADQHRCEVDVDCILEPSPC-PQRCVNTGGFCFCHYPNYDLVDGE--CVEPVDPC--FR | 371 |
| DB | 1272 | SEDMTCTVDVNECDLNPICLSGTCENTKGSFICHCDMGYSKKKGTGCTD--INECEIGA | 1330 |
| QY | 372 | ANCEYQCQPLNOT--SYLCVCAEGFA-----PIPEPHRCQMFQNOTACPADCPN | 420 |
| DB | 1331 | HNCGRHAVCTNTAGSFKSCSPGWIGDGKTKTDLDECSNGTHWCSOH-----ADC-KN | 1382 |
| QY | 421 | TOAS--CECEGYILDDGFICTDIDEC-EN-----GGF----- | 450 |
| DB | 1383 | TMGSYRCLCKDGY-TGDGFTCTDLDECSNLNLCGNGQCLNAPGGYRCEDMGFVPSADG | 1441 |
| QY | 451 | -----CS-----GVCHNLPGTFECICGPDLSALARHIG-----TDCDSK | 484 |
| DB | 1442 | KACEDIDECLEPNICVFGTCHNLPLFRCECEIGYELDRSGNCTDVNECLDPTTTCISGN | 1501 |
| QY | 485 | VDGDSGSGSEPPPTPGS-----TLTPPAVGLV | 513 |
| DB | 1502 | CVN-----TPGSYTCDCSPDFELNTRVGCV | 1527 |
| RESULT 11 | | | |
| FBL2_MOUSE STANDARD; PRT; 1221 AA. | | | |
| AC | P37889; Q9WU12; | | |
| DT | 01-OCT-1994 (Rel. 30, Created) | | |
| DT | 01-OCT-1994 (Rel. 30, Last sequence update) | | |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) | | |
| DE | FIBULIN-2 PRECURSOR. | | |
| GN | FBLN2. | | |
| OS | Mus musculus (Mouse). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| OX | NCBI_TaxID=10090; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A., AND SEQUENCE OF 27-35. | | |
| RC | TISSUE=Fibroblast; | | |
| RX | MEDLINE=94064787; PubMed=8245130; | | |
| RA | Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-L.; | | |
| RT | "Structure and expression of fibulin-2, a novel extracellular matrix | | |
| RT | protein with multiple EGF-like repeats and consensus motifs for | | |
| RT | calcium binding."; | | |
| RL | J. Cell Biol. 123:1269-1277(1993). | | |
| RN | [2] | | |
| RP | SEQUENCE FROM N.A. | | |
| RX | MEDLINE=99337686; PubMed=10406956; | | |
| RA | Graessel S., Sicot F.-X., Gotta S., Chu M.-L.; | | |
| RT | "Mouse fibulin-2 gene. Complete exon-intron organization and promoter | | |
| RT | characterization."; | | |
| RL | Eur. J. Biochem. 263:471-477(1999). | | |
| CC | -!- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS | | |
| CC | -!- CALCIUM DEPENDENT. | | |
| CC | -!- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED. | | |
| CC | -!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX. | | |
| CC | -!- ALTERNATIVE PRODUCTS: AT LEAST TWO FORMS OF FIBULIN-2 DUE TO THE | | |
| CC | PRESENCE OR ABSENCE OF A SINGLE EGF-LIKE (3) REPEAT MAY ARISE BY | | |
| CC | ALTERNATIVE SPLICING. | | |
| CC | -!- TISSUE SPECIFICITY: COMPONENT OF BOTH BASEMENT MEMBRANES AND OTHER | | |
| CC | CONNECTIVE TISSUES. | | |
| CC | -!- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS. | | |
| CC | -!- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS. | | |

| | | | | |
|----|----------|------|------|---|
| FT | SIGNAL | 1 | 26 | |
| FT | CHAIN | 27 | 1221 | FBULIN-2. |
| FT | DOMAIN | 27 | 434 | N. |
| FT | DOMAIN | 27 | 176 | SUBDOMAIN NA (CYS-RICH). |
| FT | DOMAIN | 177 | 434 | SUBDOMAIN NB (CYS-FREE). |
| FT | DOMAIN | 435 | 477 | ANAPHYLATOXIN-LIKE 1. |
| FT | DOMAIN | 478 | 510 | ANAPHYLATOXIN-LIKE 2. |
| FT | DOMAIN | 511 | 543 | ANAPHYLATOXIN-LIKE 3. |
| FT | DOMAIN | 594 | 635 | EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL). |
| FT | DOMAIN | 669 | 708 | EGF-LIKE 2. |
| FT | DOMAIN | 709 | 755 | EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL). |
| FT | DOMAIN | 756 | 800 | EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL). |
| FT | DOMAIN | 801 | 846 | EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL). |
| FT | DOMAIN | 847 | 894 | EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL). |
| FT | DOMAIN | 895 | 937 | EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL). |
| FT | DOMAIN | 938 | 979 | EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL). |
| FT | DOMAIN | 980 | 1018 | EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL). |
| FT | DOMAIN | 1019 | 1061 | EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL). |
| FT | DOMAIN | 1062 | 1106 | EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL). |
| FT | DOMAIN | 1111 | 1221 | DOMAIN III. |
| FT | SITE | 421 | 423 | CELL ATTACHMENT SITE (POTENTIAL). |
| FT | DISULFID | 435 | 462 | BY SIMILARITY. |
| FT | DISULFID | 436 | 469 | BY SIMILARITY. |
| FT | DISULFID | 449 | 470 | BY SIMILARITY. |
| FT | DISULFID | 479 | 508 | BY SIMILARITY. |
| FT | DISULFID | 492 | 509 | BY SIMILARITY. |
| FT | DISULFID | 511 | 535 | BY SIMILARITY. |
| FT | DISULFID | 512 | 542 | BY SIMILARITY. |
| FT | DISULFID | 525 | 543 | BY SIMILARITY. |

| | | | | |
|----|-----|--|----------------|-----|
| Qy | 154 | EQQCEVKAD-GFLCEHFHPATC-----RPLAV-----EPGAAAAYS----- | IT | 192 |
| Db | 521 | EQQCESNPNLIGPCN-HVMLSCGEGEPLIVPEVRPPPEAPRRVSEMEMASREALS | 579 | |
| Qy | 193 | YGT-----PFAARGADFO-----ALP-----VGSSAAVAPLGLQLM----- | CITAPPG | 229 |
| Db | 580 | LCTAEPLPNSLPGDDQDECLMLPGLCHLICINTVSGRCACPPGFELQDGRGTCRDPGG | 639 | |
| Qy | 230 | AVQGHWARE-APGAWDCSV-----ENGGCEHACNAIPGAPRCQC PAGAA | 272 | |
| Db | 640 | APQLDTARESAPRESAQVSPNTIPLVPQPNCTKONGPCRCQVRCVVGDTAMCSCPPGYA | 699 | |
| Qy | 273 | LOADGRSC-----TASATQSCNDLCERHFCVPNDPQGSYSC----- | MCETGTRLAADQHR | 322 |
| Db | 700 | IMADGVSCEDQDECLMTHDCS--WKQFQY---NTLGSYCVNHTVLCAGTYLNA-HRK | 753 | |
| Qy | 323 | CEDYDDC1LLEPSPC--PORCVNTQGGFECH-----CYPNTYDLVGGCEVPVDPCFRA-- | NC | 374 |
| Db | 754 | CVDINECVTDLHTCTRAEHCVNTPGSFQCKYALTCBPGYVLTDBECTD-VDECVTGTNHC | 812 | |
| Qy | 375 | E--YQCQPLNQTSYLCV-----CAEGFAPIPH-----EPHRQCMFCNQTCACPA | 415 | |
| Db | 813 | QAGFSQCN-TKGSFYQARQRCMDGFLQDPGEGNCVDINECTSLLEPCRSFGSCINTVGSY | 871 | |
| Qy | 416 | DCDPNTQASCECPBGY-ILDDGFICTDIDECENGFCSG---VCHNLPGTFBICGP | 468 | |
| Db | 872 | TCQRNPLV---CGRGYHANEGSECVDNCECTGVHRCBEGGOLCYNLPQSYRCDCPK | 925 | |

Best Local Similarity 29.5%; Pred. No. 1.9e-16;

Best Local Similarity 29.5%; Pred. No. 1.9e-16;

Matches 123; Conservative 39; Mismatches 141;

154 EEQQCEVKAD-GFLCEFHFPATC-----RPLAV-----EPGAA

[illegible]

b 521 EGQSCESNPNLGYPCN-HVMLSCCEGEEPLIVFEVRRPPEPEAA

193 YGT---PFAARGADFQ---ALP-----VGSSAAVAPL

[illegible]

580 LGTEAELPNSLPGDDQDECLMLPGELCQHLCINTVGSYRCACFFD

| Case | Age | Sex | Occupation | Duration of symptoms | Location of symptoms | Severity of symptoms | Response to treatment | Outcome |
|------|-----|--------|---------------|----------------------|----------------------|----------------------|-----------------------|----------|
| 1 | 45 | Male | Farmer | 10 years | Right arm | Severe | Partial | Recovery |
| 2 | 52 | Female | Teacher | 5 years | Left leg | Moderate | Complete | Recovery |
| 3 | 60 | Male | Engineer | 3 years | Right leg | Severe | Partial | Recovery |
| 4 | 58 | Female | Homemaker | 8 years | Left arm | Moderate | Complete | Recovery |
| 5 | 48 | Male | Doctor | 12 years | Right leg | Severe | Partial | Recovery |
| 6 | 55 | Female | Nurse | 7 years | Left arm | Moderate | Complete | Recovery |
| 7 | 62 | Male | Retired | 4 years | Right arm | Severe | Partial | Recovery |
| 8 | 50 | Female | Businesswoman | 6 years | Left leg | Moderate | Complete | Recovery |
| 9 | 53 | Male | Engineer | 9 years | Right leg | Severe | Partial | Recovery |
| 10 | 57 | Female | Homemaker | 11 years | Left arm | Moderate | Complete | Recovery |

Y 230 AVQGHWARE-APGAWDCSV-----ENGGEHAC

640 300 200 100 0

040 APQLDTARESA PRSESAQVSPNTIPLPVPQPNTCKDNGFCRQVC

273 10XDCBSC - - - - EACATACCNDI CEUEECVBNBDDDCBVCSC -

2/3 LQADGR3C-----IASAIQSCNDLCEHFCVFNFDQPGSISC-----

```

700 TMADGVSCFDDDECT.MGTHDCS--WKQECV--NTIGSEFYCVNHF

```

[illegible]

323 CEDVNDCTLEPSPC--PORCVNTOGGECH---CYPNYDI.VDG

| | | |
|-----|--|--|
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |
| 5 | | |
| 6 | | |
| 7 | | |
| 8 | | |
| 9 | | |
| 10 | | |
| 11 | | |
| 12 | | |
| 13 | | |
| 14 | | |
| 15 | | |
| 16 | | |
| 17 | | |
| 18 | | |
| 19 | | |
| 20 | | |
| 21 | | |
| 22 | | |
| 23 | | |
| 24 | | |
| 25 | | |
| 26 | | |
| 27 | | |
| 28 | | |
| 29 | | |
| 30 | | |
| 31 | | |
| 32 | | |
| 33 | | |
| 34 | | |
| 35 | | |
| 36 | | |
| 37 | | |
| 38 | | |
| 39 | | |
| 40 | | |
| 41 | | |
| 42 | | |
| 43 | | |
| 44 | | |
| 45 | | |
| 46 | | |
| 47 | | |
| 48 | | |
| 49 | | |
| 50 | | |
| 51 | | |
| 52 | | |
| 53 | | |
| 54 | | |
| 55 | | |
| 56 | | |
| 57 | | |
| 58 | | |
| 59 | | |
| 60 | | |
| 61 | | |
| 62 | | |
| 63 | | |
| 64 | | |
| 65 | | |
| 66 | | |
| 67 | | |
| 68 | | |
| 69 | | |
| 70 | | |
| 71 | | |
| 72 | | |
| 73 | | |
| 74 | | |
| 75 | | |
| 76 | | |
| 77 | | |
| 78 | | |
| 79 | | |
| 80 | | |
| 81 | | |
| 82 | | |
| 83 | | |
| 84 | | |
| 85 | | |
| 86 | | |
| 87 | | |
| 88 | | |
| 89 | | |
| 90 | | |
| 91 | | |
| 92 | | |
| 93 | | |
| 94 | | |
| 95 | | |
| 96 | | |
| 97 | | |
| 98 | | |
| 99 | | |
| 100 | | |

754 CVDINECVTDLHTCTRAEHCVNTPGSEOCYKAI.TCEPQYVI.TDGE

[illegible]

375 E--YOCOPLNOTSYLCV---CAEGFAP1PH-----

| | | |
|--|----|--|
| | — | |
| | — | |
| | .. | |
| | — | |
| | — | |
| | .. | |
| | — | |
| | .. | |
| | — | |
| | .. | |
| | .. | |

813 OAGFSCON-TKGSFYCOARORCMDGFELODPEGNCVDINECTSLI

.....

416 DCDPNT0ASCCEPEGY-ILDDGFICTDIDECENGGFCSG---VOC

— — — — —

b 872 TCQRNPLV- - -CGRGYHANEEGSECVDVNECETGVHRCGEGQLC

```

RESULT 12
MTN2_HUMAN
ID MTN2_HUMAN STANDARD; PRT; 956 AA.
AC O00339; O9NSZ1;
DT 30-MAY-2000 (Rel. 39, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DE MATRILIN-2 PRECURSOR.
GN MATN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP PUBLISHED FROM N.A. (LONG AND SHORT ISOFORMS).
RX PubMed=11124542;
RA Muratoglu S., Krysan K., Balazs M., Sheng H., Zakany R., Modis L.,
RA Kiss I., Deak F.;
RT "Primary structure of human matrilin-2, chromosome location of the
RT MATN2 gene and conservation of an AT-AC intron in matrilin genes.";
RT Cytogenet. Cell Genet. 90:323-327(2000).
RL [2]
RN [3]
RP SEQUENCE OF 644-956 FROM N.A.
RX MEDLINE=97238863; PubMed=9083061;
RA Deak F., Piecha D., Bachrati C., Paulsson M., Kiss I.;
RT "Primary structure and expression of matrilin-2, the closest relative
RT of cartilage matrix protein within the von Willebrand factor type A-
RT like module superfamily.";
RL J. Biol. Chem. 272:9268-9274(1997).
RN [3]
RP SEQUENCE OF 244-956 FROM N.A.
RC TISSUE=Testis;
RA Duesterhoeft A., Lauber J., Mewes H.-W., Gassenhuber J., Wiemann S.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: INVOLVED IN MATRIX ASSEMBLY (BY SIMILARITY).
CC 1- SUBCELLULAR LOCATION: SECRETED.
CC 1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC 1- SIMILARITY: CONTAINS 10 EGF-LIKE DOMAINS.
CC 1- SIMILARITY: CONTAINS 2 VWFA DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
DR EMBL; U69263; AAC51260.2; -.
DR EMBL; AL137638; CAB70853.1; ALT_INIT.
DR HSSP; P05099; 1AQ5.
DR MIM; 602108; -.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF00008; EGF; 10.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00001; EGF_like; 9.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 9.
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS0234; VWFA; 2.
KW EGF-like domain; Signal; Glycoprotein; Repeat; Coiled coil;
KW Alternative splicing.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 956 MATRILIN-2.
FT DOMAIN 57 232 VWFA 1.
FT DOMAIN 238 278 EGF-LIKE 1.
FT DOMAIN 279 319 EGF-LIKE 2.

```

```

FT DOMAIN 320 360 EGF-LIKE 3.
FT DOMAIN 361 401 EGF-LIKE 4.
FT DOMAIN 402 442 EGF-LIKE 5.
FT DOMAIN 443 483 EGF-LIKE 6.
FT DOMAIN 484 524 EGF-LIKE 7.
FT DOMAIN 525 565 EGF-LIKE 8.
FT DOMAIN 566 606 EGF-LIKE 9.
FT DOMAIN 607 647 EGF-LIKE 10.
FT DOMAIN 655 830 VWFA 2.
FT DOMAIN 917 955 COILED COIL (POTENTIAL).
FT DISULFID 242 253 BY SIMILARITY.
FT DISULFID 249 262 BY SIMILARITY.
FT DISULFID 264 277 BY SIMILARITY.
FT DISULFID 283 294 BY SIMILARITY.
FT DISULFID 290 303 BY SIMILARITY.
FT DISULFID 305 318 BY SIMILARITY.
FT DISULFID 324 335 BY SIMILARITY.
FT DISULFID 331 344 BY SIMILARITY.
FT DISULFID 346 359 BY SIMILARITY.
FT DISULFID 365 376 BY SIMILARITY.
FT DISULFID 372 385 BY SIMILARITY.
FT DISULFID 387 400 BY SIMILARITY.
FT DISULFID 406 417 BY SIMILARITY.
FT DISULFID 413 426 BY SIMILARITY.
FT DISULFID 428 441 BY SIMILARITY.
FT DISULFID 447 458 BY SIMILARITY.
FT DISULFID 454 467 BY SIMILARITY.
FT DISULFID 469 482 BY SIMILARITY.
FT DISULFID 488 499 BY SIMILARITY.
FT DISULFID 495 508 BY SIMILARITY.
FT DISULFID 510 523 BY SIMILARITY.
FT DISULFID 529 540 BY SIMILARITY.
FT DISULFID 536 549 BY SIMILARITY.
FT DISULFID 551 564 BY SIMILARITY.
FT DISULFID 570 581 BY SIMILARITY.
FT DISULFID 577 590 BY SIMILARITY.
FT DISULFID 592 605 BY SIMILARITY.
FT DISULFID 611 622 BY SIMILARITY.
FT DISULFID 618 631 BY SIMILARITY.
FT DISULFID 633 646 BY SIMILARITY.
FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 890 890 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPIC 861 879 MISSING (IN SHORT ISOFORM).
FT CONFLICT 356 356 K -> E (IN REF. 3).
FT CONFLICT 594 594 R -> E (IN REF. 3).
FT CONFLICT 644 644 R -> G (IN REF. 2).
FT CONFLICT 755 755 F -> L (IN REF. 3).
FT CONFLICT 935 935 L -> F (IN REF. 2).
SQ SEQUENCE 956 AA; 106840 MW; 826B7F347178FC80 CRC64;

Query Match 12.2%; Score 356; DB 1; Length 956;
Best Local Similarity 31.6%; Pred. No. 1.8e-16;
Matches 86; Conservative 40; Mismatches 102; Indels 44; Gaps 15;

QY 245 CSVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTA-SATQSCNDLCEHFCVNPDPQ 303
Db 283 CAMEDHNCEQLCVNVPGSFVCQYSGYALAEGRKRVAVDYCASNHGCEHCV-NAD-- 339
QY 304 GSYSCMCETGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFCHCYPNYDL-VDGEC 362
Db 340 GSYLCQCHGEFALNPDKTKTKIDYCASSNHGCEHCVNTDDSYSCHLKGFTLNPDKKT 399
QY 363 VEPVDPCT--RANCEYQCQPLNQTSYLCVCAEGFAPIPH-----EPHRCQMF 408
Db 400 CRRINYCALNKPGEHECVNMEE-SYCRCHRGVTLDPNGTKSRVDRHCAQDHCCEQLC 458
QY 409 NQTACPADCDNTQAS--CECPGEYILDDGF-ICTDIDEC----ENGFCSGVCHNLPTFF 462
Db 459 -----LNTEDSFVCQSEGLINEDLKTCSRVDYCLLSDHG--CEYSCVNDRSF 506
QY 463 ECICGPDALARHGTDCDSKVDG---GDSG 491

```

| | | | |
|------------|--|------------------------------------|---|
| Db | 507 | ACQC-PEGHVLRSQKTC--AKLDSALGDBG 535 | |
| RESULT 13 | | | |
| NTG4_MOUSE | | | |
| ID | NTG4_MOUSE | STANDARD; | PRT; 1964 AA. |
| AC | P31695; Q62389; | | |
| DT | 01-JUL-1993 (Rel. 26, Created) | | |
| DT | 01-JUL-1997 (Rel. 35, Last sequence update) | | |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) | | |
| DE | NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 4 PRECURSOR (TRANSFORMING | | |
| DE | PROTEIN INT-3). | | |
| GN | NOTCH4 OR INT3 OR INT-3. | | |
| OS | Mus musculus (Mouse). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| OX | NCBI_TaxID=10090; | | |
| [1] | | | |
| RN | SEQUENCE FROM N.A. | | |
| RX | MEDLINE=92194507; PubMed=1312643; | | |
| RA | Robbins J., Blondel B.J., Gallahan D., Callahan R.; | | |
| RT | "Mouse mammary tumor gene int-3: a member of the notch gene family | | |
| RT | transforms mammary epithelial cells.;" | | |
| RL | J. Virol. 66:2594-2599(1992). | | |
| [2] | | | |
| RN | REVISIONS, SEQUENCE FROM N.A. | | |
| RX | MEDLINE=97294599; PubMed=9150355; | | |
| RA | Gallahan D., Callahan R.; | | |
| RT | "The mouse mammary tumor associated gene INT3 is a unique member of | | |
| RT | the NOTCH gene family (NOTCH4).;" | | |
| RL | Oncogene 14:1883-1890(1997). | | |
| [3] | | | |
| RN | SEQUENCE FROM N.A. | | |
| RC | TISSUE=Lung, and Testis; | | |
| RX | MEDLINE=96281668; PubMed=8681805; | | |
| RA | Uyttendaele H., Marazzi G., Wu G., Yan Q., Sassoon D., Kitajewski J.; | | |
| RT | "Notch4/int-3, a mammary proto-oncogene, is an endothelial | | |
| RT | cell-specific mammalian Notch gene.;" | | |
| RL | Development 122:2251-2259(1996) | | |
| CC | -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. | | |
| CC | -1- DISEASE: ACTIVATED INT-3 TRANSFORMS MAMMARY EPITHELIAL CELLS. | | |
| CC | -1- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS. | | |
| CC | -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS. | | |
| CC | -1- SIMILARITY: CONTAINS 5 ANK REPEATS. | | |
| CC | ----- | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | |
| CC | use by non-profit institutions as long as its content is in no way | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | |
| CC | or send an email to license@isb-sib.ch). | | |
| CC | ----- | | |
| DR | EMBL; M80456; AAB38377.1; - | | |
| DR | EMBL; U43691; AAC52630.1; - | | |
| DR | PIR; A38072; TWMV73. | | |
| DR | HSSP; P00740; IIXA. | | |
| DR | MGD; MGI:107471; Notch4. | | |
| DR | InterPro; IPR002110; ANK. | | |
| DR | InterPro; IPR000152; Asx_hydroxyl. | | |
| DR | InterPro; IPR000561; EGF-like. | | |
| DR | InterPro; IPR000742; EGF_2. | | |
| DR | InterPro; IPR001881; EGF_Ca. | | |
| DR | InterPro; IPR001438; EGF_II. | | |
| DR | InterPro; IPR000800; Notch. | | |
| DR | Pfam; PF00023; ank; 6 | | |
| DR | Pfam; PF00008; EGF; 27 | | |
| DR | Pfam; PF00066; notch; 2 | | |
| DR | PRINTS; PR00010; EGFBLD. | | |
| DR | SMART; SM00248; ANK; 5 | | |
| DR | SMART; SM00179; EGF_CA; 11 | | |
| DR | SMART; SM00001; EGF_like; 15 | | |
| DR | SMART; SM00004; NL; 2 | | |
| DR | | | |
| DR | PROSITE; PS50088; ANK_REPEAT; 5. | | |
| DR | PROSITE; PS00297; ANK_REPEAT_REGION; 1. | | |
| DR | PROSITE; PS00010; ASX_HYDROXYL; 11. | | |
| DR | PROSITE; PS00022; EGF_1; 28. | | |
| DR | PROSITE; PS01186; EGF_2; 21. | | |
| DR | PROSITE; PS01187; EGF_CA; 9. | | |
| KW | Differentiation; Neurogenesis; Repeat; EGF-like domain; Transmembrane; | | |
| KW | Glycoprotein; Proto-oncogene; ANK repeat; Signal. | | |
| FT | SIGNAL 1 20 | | POTENTIAL. |
| FT | CHAIN 21 1964 | | NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 4. |
| FT | DOMAIN 21 1443 | | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM 1444 1464 | | POTENTIAL. |
| FT | DOMAIN 1465 1964 | | CYTOPLASMIC (POTENTIAL). |
| FT | DOMAIN 21 60 | | EGF-LIKE 1. |
| FT | DOMAIN 61 112 | | EGF-LIKE 2. |
| FT | DOMAIN 115 152 | | EGF-LIKE 3. |
| FT | DOMAIN 153 189 | | EGF-LIKE 4. |
| FT | DOMAIN 191 229 | | EGF-LIKE 5. |
| FT | DOMAIN 231 271 | | EGF-LIKE 6. |
| FT | DOMAIN 273 309 | | EGF-LIKE 7. |
| FT | DOMAIN 311 350 | | EGF-LIKE 8. |
| FT | DOMAIN 352 388 | | EGF-LIKE 9. |
| FT | DOMAIN 389 427 | | EGF-LIKE 10. |
| FT | DOMAIN 429 470 | | EGF-LIKE 11. |
| FT | DOMAIN 472 508 | | EGF-LIKE 12. |
| FT | DOMAIN 510 546 | | EGF-LIKE 13. |
| FT | DOMAIN 548 584 | | EGF-LIKE 14. |
| FT | DOMAIN 586 622 | | EGF-LIKE 15. |
| FT | DOMAIN 622 656 | | EGF-LIKE 16. |
| FT | DOMAIN 658 686 | | EGF-LIKE 17. |
| FT | DOMAIN 688 724 | | EGF-LIKE 18. |
| FT | DOMAIN 726 762 | | EGF-LIKE 19. |
| FT | DOMAIN 764 800 | | EGF-LIKE 20. |
| FT | DOMAIN 803 839 | | EGF-LIKE 21. |
| FT | DOMAIN 841 877 | | EGF-LIKE 22. |
| FT | DOMAIN 878 924 | | EGF-LIKE 23. |
| FT | DOMAIN 926 962 | | EGF-LIKE 24. |
| FT | DOMAIN 964 1000 | | EGF-LIKE 25. |
| FT | DOMAIN 1002 1040 | | EGF-LIKE 26. |
| FT | DOMAIN 1042 1081 | | EGF-LIKE 27. |
| FT | DOMAIN 1083 1122 | | EGF-LIKE 28. |
| FT | DOMAIN 1126 1167 | | EGF-LIKE 29. |
| FT | REPEAT 1168 1208 | | LIN/NOTCH 1. |
| FT | REPEAT 1209 1242 | | LIN/NOTCH 2. |
| FT | REPEAT 1243 1282 | | LIN/NOTCH 3. |
| FT | REPEAT 1628 1657 | | ANK 1. |
| FT | REPEAT 1661 1691 | | ANK 2. |
| FT | REPEAT 1695 1724 | | ANK 3. |
| FT | REPEAT 1728 1757 | | ANK 4. |
| FT | REPEAT 1761 1790 | | ANK 5. |
| FT | DISULFID 25 38 | | BY SIMILARITY. |
| FT | DISULFID 32 48 | | BY SIMILARITY. |
| FT | DISULFID 50 59 | | BY SIMILARITY. |
| FT | DISULFID 65 77 | | BY SIMILARITY. |
| FT | DISULFID 71 100 | | BY SIMILARITY. |
| FT | DISULFID 102 111 | | BY SIMILARITY. |
| FT | DISULFID 119 130 | | BY SIMILARITY. |
| FT | DISULFID 124 140 | | BY SIMILARITY. |
| FT | DISULFID 142 151 | | BY SIMILARITY. |
| FT | DISULFID 157 168 | | BY SIMILARITY. |
| FT | DISULFID 162 177 | | BY SIMILARITY. |
| FT | DISULFID 179 188 | | BY SIMILARITY. |
| FT | DISULFID 195 208 | | BY SIMILARITY. |
| FT | DISULFID 202 217 | | BY SIMILARITY. |
| FT | DISULFID 219 228 | | BY SIMILARITY. |
| FT | DISULFID 235 246 | | BY SIMILARITY. |
| FT | DISULFID 240 259 | | BY SIMILARITY. |
| FT | DISULFID 261 270 | | BY SIMILARITY. |
| FT | DISULFID 277 288 | | BY SIMILARITY. |
| FT | DISULFID 282 297 | | BY SIMILARITY. |
| FT | DISULFID 299 308 | | BY SIMILARITY. |
| FT | DISULFID 315 329 | | BY SIMILARITY. |
| FT | DISULFID 323 338 | | BY SIMILARITY. |

Search completed: May 8, 2002, 12:45:19
Job time: 355 sec

